

This electronic thesis or dissertation has been downloaded from the King's Research Portal at <https://kclpure.kcl.ac.uk/portal/>



Immunohistochemical and Genomic Analysis of Ductal Carcinoma in Situ of the Human Breast

Brown, John Peter

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT



Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. <https://creativecommons.org/licenses/by-nc-nd/4.0/>

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

**Immunohistochemical and
Genomic Analysis of Ductal
Carcinoma in Situ of the Human
Breast**

**King's College London
2016**

John Peter Brown

PhD

Abstract

Ductal Carcinoma in Situ (DCIS) is a non obligate precursor of invasive breast cancer. Due mainly to improved screening methods the detection of DCIS has risen over the last twenty years. The inability to reliably predict progression to invasive breast cancer results in the possible overtreatment of what can be regarded as a non-life threatening condition. Current treatment options remain controversial with no consensus upon the choice of mastectomy or breast conserving surgery with or without radiotherapy or adjuvant hormone therapies. The natural progression and molecular pathology of DCIS also still remains poorly understood.

Examination of DCIS cases with known clinical outcome has been carried out using traditional pathological criteria, immunohistochemistry and genomic analysis. Tissue microarrays have been constructed and stained with a panel of antibody markers demonstrating the presence of molecular subtypes of DCIS similar to those found in invasive breast disease, although at different frequencies.

Molecular inversion probe arrays have revealed a complex heterogeneity exists in early stages of breast cancer that may be drivers of invasion. Comparison with invasive breast disease and with DCIS associated with invasive breast disease reveals a highly complex system, where progression may rely on genomic changes already established within pure DCIS or those that occur during the DCIS phase of tumourigenesis.

ABSTRACT	2
ACKNOWLEDGEMENTS	15
ABBREVIATIONS AND GLOSSARY OF TERMS	16
CHAPTER 1. INTRODUCTION	18
1.1 THE NORMAL BREAST AND DUCTAL CARCINOMA IN SITU	18
1.2 EPIDEMIOLOGY OF DUCTAL CARCINOMA IN SITU.....	19
1.2.1 <i>Incidence of ductal carcinoma in situ (DCIS)</i>	19
1.2.2 <i>Ductal Carcinoma in Situ and Gender</i>	21
1.2.3 <i>Risk Factors of Ductal Carcinoma in Situ</i>	21
1.2.4 <i>Risk of Ductal Carcinoma in Situ Progressing to Invasive Disease</i>	22
1.3 DETECTION OF DUCTAL CARCINOMA IN SITU.....	23
1.2.1 <i>Mammography</i>	23
1.2.2 <i>Ultrasound</i>	24
1.2.3 <i>Magnetic Resonance Imaging (MRI)</i>	25
1.3 HISTOLOGY OF DUCTAL CARCINOMA IN SITU.....	25
1.3.1 <i>Histological Diagnosis of Ductal Carcinoma in Situ</i>	25
1.3.2 <i>Usual Epithelial Hyperplasia</i>	26
1.3.3 <i>Columnar Cell Lesions (CCLs) including Flat Epithelial Atypia (FEA)</i>	27
1.3.4 <i>Atypical Ductal Hyperplasia (ADH)</i>	28
1.3.5 <i>Lobular Carcinoma In Situ, Atypical Lobular Hyperplasia, Lobular Neoplasia</i>	29
1.3.6 <i>Microinvasive Carcinoma</i>	30
1.3.7 <i>Histological Variants of Ductal Carcinoma in Situ</i>	31
1.3.8 <i>Rare Histological Forms of Ductal Carcinoma in Situ</i>	34
1.4 CLASSIFICATION SYSTEMS FOR DUCTAL CARCINOMA IN SITU	35
1.4.1 <i>Growth Patterns for Ductal Carcinoma in Situ</i>	35
1.4.2 <i>High Grade, Intermediate Grade and Low Grade Ductal Carcinoma in Situ</i>	35
1.4.3 <i>The Van Nuys Grading System for DCIS and the Van Nuys Prognostic Index (VNPI)</i>	36
1.4.4 <i>The Tavassoli Grading/Classification System</i>	37

1.4.5	<i>R Holland and the European Pathologists Working Group Classification System</i>	37
1.4.6	<i>Summary of Classification of DCIS</i>	38
1.5	TREATMENT OF DUCTAL CARCINOMA IN SITU	39
1.5.1	<i>Surgical Procedures and Options for Ductal Carcinoma in Situ</i>	39
1.5.2	<i>Radiotherapy for Ductal Carcinoma in Situ</i>	39
1.5.3	<i>Hormone Therapy for Ductal Carcinoma in Situ</i>	40
1.5.4	<i>Chemotherapy for Ductal Carcinoma in Situ</i>	41
1.6	TECHNIQUES FOR THE DIAGNOSIS AND INVESTIGATION OF PROGNOSIS IN DCIS AND INVASIVE BREAST CANCER	
	41	
1.6.1	<i>Sampling Techniques</i>	41
1.6.2	<i>Fixation and Processing of Biopsies</i>	42
1.6.3	<i>Immunohistochemistry in Assessment and Prognosis of Ductal Carcinoma in Situ</i>	44
1.6.4	<i>Tissue Microarrays (TMAs)</i>	45
1.6.5	<i>Genetic Aberrations & Molecular Inversion Probe Arrays for Genomic Profiling</i>	46
1.7	MOLECULAR PORTRAITS OF BREAST CANCER	50
1.7.1	<i>Genetic Subgroups of Invasive Breast Cancer</i>	50
1.7.2	<i>IHC as a Surrogate to Genomic Profiling</i>	53
1.8	CONCLUSION	55
1.9	HYPOTHESIS	56
 CHAPTER 2. CONSTRUCTION OF TISSUE MICROARRAYS (TMAS) FOR IMMUNOHISTOCHEMICAL		
EVALUATION OF DUCTAL CARCINOMA IN SITU		
57		
2.1	INTRODUCTION	57
2.2	MATERIALS AND METHODS	57
2.3	RESULTS	61
2.3.1	<i>TMA Constructs</i>	61
2.3.2	<i>Scanning</i>	61
2.4	DISCUSSION	61
2.4.1	<i>Tissue Selection Criteria</i>	61

2.4.2	<i>TMA cores size, number and distribution within a TMA block</i>	62
CHAPTER 3. IMMUNOHISTOCHEMICAL STUDIES ON DUCTAL CARCINOMA IN SITU..... 65		
3.1	INTRODUCTION	65
3.2	MATERIAL AND METHODS.....	65
3.2.1	<i>Correlation of Immunohistochemical Results to Patient Data and Outcome.....</i>	65
3.2.2	<i>Histological Assessment of Necrosis and Inflammatory Response.</i>	66
3.2.3	<i>Immunohistochemical Experiments.....</i>	68
3.3	HORMONE RECEPTORS	69
3.3.1	<i>Oestrogen Receptor Structure and Function</i>	69
3.3.2	<i>Oestrogen Receptor and Breast Cancer.....</i>	70
3.3.3	<i>The Oestrogen Receptor and Ductal Carcinoma in Situ.....</i>	71
3.3.4	<i>Progesterone Receptor</i>	71
3.4	EPIDERMAL GROWTH FACTOR RECEPTORS.....	72
3.4.1	<i>Epidermal Growth Factor Receptor Structure and Function.....</i>	72
3.5	PROLIFERATION MARKERS.....	74
3.5.1	<i>Introduction.....</i>	74
3.5.2	<i>Ki-67.....</i>	74
3.5.3	<i>Mini Chromosome Maintenance Proteins</i>	78
3.6	BASAL MARKERS	81
3.6.1	<i>Cytokeratins 5, 6 and 14.....</i>	81
3.7	SCORING OF ANTIBODY MARKERS IN IMMUNOHISTOCHEMICAL STUDIES.....	82
3.7.1	<i>Analysis of differing IHC Scores.</i>	88
3.7.2	<i>Inter Observer Variability in Scoring Immunohistochemistry</i>	89
3.8	STATISTICAL ANALYSIS OF IMMUNOHISTOCHEMICAL SCORING	90
3.9	IMMUNOHISTOCHEMISTRY RESULTS	90
3.9.1	<i>Immunohistochemical Scoring</i>	90
3.9.2	<i>Immunohistochemical Images.....</i>	93
3.9.3	<i>Non Parametric Tests- Univariate Paired analysis:</i>	96

3.9.4	<i>Univariate Analysis of Hormone Receptor Immunohistochemistry</i>	97
3.9.5	<i>Univariate Analysis for HER2 Immunohistochemistry</i>	102
3.9.6	<i>Univariate Analysis for Proliferation Marker Immunohistochemistry</i>	104
3.9.7	<i>Univariate Analysis for Cytokeratin Immunohistochemistry</i>	107
3.9.8	<i>Univariate Analysis and Recurrence of DCIS and/or Invasive Breast Disease</i>	110
3.9.9	<i>Histological Grade in DCIS</i>	124
3.9.10	<i>Assignment of Molecular Subtypes to DCIS</i>	132
3.10	IMMUNOPHENOTYPIC SUBTYPES OF DCIS AND RECURRENCE	133
3.11	DISCUSSION	134
CHAPTER 4.	DCIS ASSOCIATED WITH INVASIVE BREAST CARCINOMA	139
4.1	INTRODUCTION	139
4.2	MATERIALS AND METHODS	139
4.3	RESULTS	141
4.3.1	<i>Individual IHC Antibody Staining Concordance between Invasive Tumour and DCIS in the Same Case:</i>	141
4.3.2	<i>Assignment of Molecular Subgroups</i>	142
4.4	DISCUSSION	144
CHAPTER 5.	GENOMIC STUDIES ON DCIS	147
5.1	INTRODUCTION	147
5.2	GENETIC ABERRATIONS OF DCIS AND INVASIVE BREAST DISEASE – LITERATURE REVIEW	148
5.3	MATERIALS AND METHODS	149
5.4	HYPOTHESIS FOR MIP ARRAYS STUDIES IN DCIS AND INVASIVE BREAST CARCINOMA	150
5.4.1	<i>DNA Extraction</i>	151
5.4.2	<i>MIP Array</i>	152
5.4.3	<i>Statistical methods</i>	152
5.4.4	<i>Analyses of Raw Data Tables</i>	153
5.5.5	<i>Analysis of MIP array Data Using “PANTHER” Classification System</i>	154
5.5.6	<i>Analysis of DCIS and Invasive Breast Disease Subgroups.</i>	155

5.6	RESULTS.....	156
5.6.1	<i>DNA Extraction for MIP Arrays</i>	156
5.6.2	<i>MIP Array Maps.....</i>	156
5.6.3	<i>Copy Number Aberrations for Oestrogen Receptor Positive DCIS Compared to Non Oestrogen Receptor DCIS.....</i>	163
5.6.4	<i>Copy Number Aberrations for Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease</i>	176
5.6.5	<i>Copy Number Aberrations for Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Positive Invasive Breast Disease</i>	193
5.6.6	<i>Copy Number Aberrations for All Oestrogen Receptor Positive DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease</i>	209
5.6.7	<i>Copy Number Aberrations for Triple Negative DCIS Compared to Non Triple negative DCIS</i> <i>220</i>	
5.6.8	<i>Copy Number Aberrations for Pure Triple Negative DCIS Compared to Triple Negative DCIS Associated with Invasive Breast Disease</i>	236
5.6.9	<i>Copy Number Aberrations for Triple Negative Pure DCIS Compared to Triple Negative Invasive Breast Disease</i>	251
5.6.10	<i>Copy Number Aberrations in All Triple Negative DCIS compared to Triple Negative Invasive breast disease.....</i>	267
5.6.11	<i>Copy Number Aberrations for HER2 Positive DCIS Compared to Oestrogen Receptor Positive DCIS and Triple Negative DCIS.....</i>	286
5.7	SUMMARY AND DISCUSSION OF GENOMIC DIFFERENCES BETWEEN PURE DCIS AND DCIS ASSOCIATED WITH INVASIVE BREAST DISEASE	301
5.7.1	<i>Genomic Differences found in Oestrogen Receptor Positive Pure DCIS Versus Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease</i>	301
5.7.2	<i>Genomic Differences found in Triple Negative Pure DCIS Versus Triple Negative DCIS Associated with Invasive Breast Disease</i>	310
5.7.3	<i>Observations on Genomic Analysis of Pure DCIS Compared to DCIS Associated with Invasive Breast Disease.</i>	316

5.8	SUMMARY OF GENOMIC DIFFERENCES BETWEEN DCIS AND INVASIVE BREAST DISEASE.....	317
5.8.1	<i>Genomic Difference between ER Positive DCIS and ER Positive Invasive Breast Disease.</i> <i>(See section 5.3.4).....</i>	318
5.8.2	<i>Genomic Difference between Triple Negative DCIS and Triple Negative Invasive Breast Disease.....</i>	320
5.8.3	<i>Observations on Genomic Analysis of DCIS Compared to DCIS Invasive Breast Disease.....</i>	324
5.9	SUMMARY OF GENOMIC DIFFERENCES BETWEEN PURE DCIS VERSUS INVASIVE BREAST DISEASE.....	325
5.9.1	<i>Pure ER Positive DCIS compared to ER Positive Invasive Breast Disease.....</i>	325
5.9.2	<i>Pure Triple Negative DCIS versus Triple Negative Invasive Breast Disease (See section 5.3.11). 327</i>	
5.9.3	<i>Observations of Genomic Analysis of Pure DCIS Compared to Invasive Breast Disease.....</i>	330
5.10	SUMMARY OF HER2 POSITIVE DCIS VERSUS ER POSITIVE DCIS AND TRIPLE NEGATIVE DCIS.....	331
5.10.1	<i>Observations of Genomic Analysis of Pure Her2 DCIS compared to Pure ER positive DCIS and Pure TN DCIS.....</i>	334
CHAPTER 6.	CONCLUSIONS.....	335
6.1	TISSUE MICROARRAYS	335
6.2	IMMUNOHISTOCHEMICAL STUDIES OF DCIS	335
6.3	GENOMIC STUDIES OF DCIS	336
6.4	LIMITATIONS OF THE STUDY	340
6.4.1	<i>Availability of clinical data</i>	340
6.4.2	<i>Technical Considerations</i>	341
6.4.3	<i>Genomics</i>	342
6.5	FURTHER WORK.....	343
	REFERENCES.....	344
	APPENDICES.....	360

APPENDIX 1: COPYRIGHT PERMISSIONS.....	360
APPENDIX 2: EXCEL TMA MAPS AND IHC SCORES.....	361
APPENDIX 3: TISSUE MICROARRAY PATIENT DEMOGRAPHICS.....	367
APPENDIX 4: IMMUNOHISTOCHEMISTRY PROTOCOLS.....	379
APPENDIX 5: FREQUENCY TABLES AND NON PARAMETRIC STATISTICAL ANALYSIS OF DCIS IHC	380
APPENDIX 6: DNA EXTRACTION AND PURIFICATION PROTOCOLS	393
APPENDIX 7. QUBIT DNA QUANTITIES QUALITY: DNA QUBIT READINGS FOR MIP ARRAY ANALYSIS	397
APPENDIX 8. MIP ARRAY ANALYSIS- DNA SAMPLES SENT TO AFFYMETRIX.....	405
APPENDIX 9 MIP ARRAY CHARTS	409

List of Figures and Tables

Figure 1: Cranio-caudal and medio-lateral oblique mammographic views of a case with two needle-localisation wires in place to mark an area of abnormality.	23
Figure 2: Usual epithelial hyperplasia	27
Figure 3: Flat epithelial atypia	28
Figure 4: Atypical Ductal Hyperplasia	29
Figure 5. Lobular carcinoma in-situ.....	30
Figure 6: DCIS (top right) with microinvasion (left).	31
Figure 7: Central comedo necrosis within DCIS	32
Figure 8: Solid architecture DCIS	33
Figure 9: Papillary DCIS/papillary carcinoma in situ	33
Figure 10: Cribriform DCIS.....	33
Figure 11: TMA manufacture and stylets	46
Figure 12: MIP Array Probe:	48
Figure 13: MIP array process for genomic profiling	49
Figure 14. Optical marking of disease on an H&E section.	58
Figure 15. Template for TMA composed of 2.0mm cores	59

Figure 16: DCIS in 2.0mm cores.....	60
Figure 17: TMA construct.....	60
Figure 18: Structure and Functional Domains of ER	70
Figure 19: EGFR Structure. DOI: http://dx.doi.org/10.7554/eLife.00708.003 reproduced through Creative Commons Attribution license: original image from reference (132)	73
Figure 20: Ki-67: Protein structure 1R-21-1-120 (human) A & 2AFF-1-120 (human) B.	76
Figure 21: Examples of a Hexameric Helicase Protein Structure for DNA replication.	80
Figure 22: MCM Proteins in the Cell Cycle	80
Figure 23: ER staining of DCIS in 2mm TMA core (X40, X200)	93
Figure 24: PR staining of DCIS in 2mm TMA core (X40, X200)	93
Figure 25: HER2 staining of DCIS in 2mm TMA core (X100, X200)	93
Figure 26: EGFR staining of DCIS in 2mm TMA core (X40, X100, X200)	94
Figure 27: CK14 staining of DCIS in 2mm TMA core (X40, X200).	94
Figure 28: Ki67 staining of DCIS in 2mm TMA core (X40, X200)	94
Figure 29: CISH for HER2 showing amplification of the <i>HER2</i> Gene in DCIS, 2mm Core (x40, X200).....	95
Figure 30: SMMHC IHC showing intact myoepithelial cell layer around DCIS.	140
Figure 31: Chromosomal map for a ER Positive pure DCIS Case (DCISJPB10K 066).	158
Figure 32: Chromosomal Map for a Pure DCIS Triple Negative Case (DCISJPB10K001) .	159
Figure 33: Chromosomal Map for a Pure HER2 Positive Case (DCISJPB10K070); note chromosome 17 allelic amplification.	160
Figure 34: Chromosomal map for normal, DCIS and invasive samples from an ER positive lesion.	161
Figure 35: Chromosomal Map for normal, DCIS and invasive samples from a triple negative case.	162
Figure 36: Frequency plots showing copy number aberrations between oestrogen receptor positive DCIS and non oestrogen positive DCIS (pages 165-167).....	163
Figure 37: Frequency plots showing copy number aberrations between oestrogen positive pure DCIS and oestrogen positive DCIS associated with tumour (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 177-180).	176

Figure 38: Frequency plots showing copy number aberrations between oestrogen receptor positive pure DCIS and oestrogen positive invasive breast disease (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 194-197).	193
Figure 39: Frequency plots showing copy number aberrations between oestrogen positive DCIS and oestrogen positive invasive breast disease (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 209-212).	209
Figure 40: Frequency plots for copy number aberrations for triple negative DCIS versus non triple negative DCIS (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH, pages 220-223).	220
Figure 41: frequency plots showing copy number aberrations between pure triple negative dcis and triple negative DCIS associated with invasive breast disease (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 236-238).	236
Figure 42: Frequency plots showing copy number aberrations between pure triple negative DCIS and triple negative invasive breast disease (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 249-251).	251
Figure 43: Frequency plots showing copy number aberrations between triple negative DCIS and triple negative invasive disease (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 264-266).	268
Figure 44: Frequency plots showing copy number aberrations between HER2 positive pure DCIS compared to oestrogen positive dcis and triple negative DCIS (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 279-282).	286
Figure 45: Molecular functions associated with Sc gains in ER positive DCIS associated with invasive breast diseases and ER positive pure DCIS.	304
Figure 46: Biological processes associated with Sc gains in ER positive DCIS associated with invasive breast disease and ER positive pure DCIS	305
Figure 47: Protein classes associated with ER positive DCIS associated with invasive breast disease and ER positive pure DCIS.	305
Figure 48: Molecular functions associated with total loss in ER positive DCIS associated with invasive breast disease.	306

Figure 49: Biological processes associated with total loss in ER positive DCIS associated with invasive breast disease.....	307
Figure 50: Protein classes associated with ER positive DCIS associated with invasive breast disease.	307
Figure 51: Molecular functions associated with CnLOH in ER positive pure DCIS	308
Figure 52: Biological processes associated with CnLOH in ER positive pure DCIS	309
Figure 53: Protein classes associated with CnLOH in ER positive pure DCIS.....	309
Figure 54: Molecular functions associated with the genomic gains identified in TN DCIS associated with invasive breast cancer.....	312
Figure 55: Biological processes associated with the gains identified in TN DCIS associated with invasive breast cancer.	312
Figure 56: Protein classes associated with the genomic gains identified in TN DCIS associated with invasive breast cancer.	313
Figure 57: Molecular functions associated with Sc gains for TN DCIS associated with invasive breast disease.	313
Figure 58: Biological processes associated with Sc gains for TN DCIS associated with invasive breast disease.	314
Figure 59: Protein classes associated with Sc gains in TN DCIS associated with invasive breast disease.	314
Figure 60: Molecular functions associated with total loss in TN DCIS	322
Figure 61: Biological processes associated with total loss in TN DCIS.....	322
Figure 62: Protein classes associated with total loss in TN DCIS	323

Tables:

Table 1: Surface areas of different TMA core sizes.....	63
Table 2: Patient Demographics and Histological Details.	67
Table 3: Allred Scoring Method.....	90
Table 4: IHC methods of scoring, of handling differing scores from cores and defined acceptable variation for DCIS immunostaining.	91
Table 5: Staining frequencies for immunohistochemical antibody staining.	92

Table 6: Univariate analysis p-values for immunohistochemical and histological parameters.	96
Table 7: Univariate paired analysis for oestrogen receptor immunohistochemistry	97
Table 8: Univariate paired analysis for progesterone receptor immunohistochemistry.	98
Table 9: Univariate analyses for ER and HER2 IHC: Fishers exact test p value < 0.0001 ...	99
Table 10: Univariate analysis for ER and PR IHC; Fishers exact test p value <0.0001	99
Table 11: Univariate analyses for ER and CK5 IHC; Fishers exact test p value = 0.0033..	100
Table 12: Univariate analyses for ER and CK14 IHC; Fishers exact test p value = 0.0018.	100
Table 13: Univariate analyses for ER and Ki67 (>5%) IHC; Fishers exact test p value <0.0001	100
Table 14: Univariate analysis for ER and MCM2; (>5%) IHC Fishers exact test p value = 0.0032	101
Table 15: Univariate analyses for ER and Histological Grade; ER; Chi Square p value <0.0001	101
Table 16: Univariate analyses for ER IHC and Histological Architecture; Chi square p Value < 0.0001	101
Table 17: Univariate paired analysis for HER2 immunohistochemistry.	102
Table 18: Univariate Analysis of Proliferation Marker Immunohistochemistry.....	105
Table 19: Paired univariate analysis for CK immunohistochemistry.....	109
Table 20: Recurrence data for IHC Markers and Luminal Groups	123
Table 21: Biomarker expression by cytonuclear grade of DCIS.	124
Table 22: Paired Univariate Analysis for Histological Grade.	131
Table 23: Definitions for Surrogate Molecular Subgroups for DCIS:	132
Table 24: Subgrouping of ER positive DCIS with Ki67 expression; cut-off defined at median, 5%.	133
Table 25: Comparison of IHC scores between cases of admixed DCIS and invasive carcinoma	142
Table 26: IHC Molecular Subgroups for Invasive Carcinomas and Associated DCIS.....	143

Table 27: IHC Molecular Subgroups for Invasive Carcinomas and Associated DCIS for ER Positive and Ki67 Expression	143
Table 28: Sample Selection and Case Numbers for MIP Array Analysis.	149
Table 29: Copy number analysis types for MIP arrays.	155
Table 30: MIP array cases with chromosomal maps	156
Table 31: MIP Array Comparison Groups.....	157
Table 32: Patient Demographics: RID= Research Number	369
Table 33: <i>IHC</i> scores for TMA cases with associated patient data.	372

Acknowledgements

There are so many people who have helped me in this endeavour. Firstly I would like to thank Professor Sarah Pinder whom I count as a true friend and also an excellent supervisor. She has been inspirational and also shown great patience in me. I may never be able to thank her enough. Sarah, I said I would submit before Watford FC were in the Premiership again I didn't think it would take quite so long for both to happen.

I would like to sincerely thank the following King's College London colleagues whom, due to their respective expertise, technical abilities and kindness have enabled me to produce this work: Professor Arnie Puroshotham, Professor Andrew Tutt, Dr Cheryl Gillett, Dr Patrycja Gazinska, Vadna Shah, Gursharn Hutchins and Sora Vorapon.

Dr Anita Grigoriadis and Hasan Mirza deserve medals for taking the time to explain to me (in fairly small words) the world of genetics and bioinformatics. Thanks for being so patient with me.

I would like to thank my sponsors Breakthrough Breast Cancer and The Breast Cancer Research Trust whose significant charitable donations have paid for this research. I would like to thank the Guy's and St Thomas's Breast Tissue and Data Bank staff for all their help.

This has been a long pathway and there are too many people to list here, apologies if you are missing but the following people most definitely deserve a mention: Louise Bell, Pierre Lao-Sirieix, Lisa Happerfield, Jill Davis, Ben Onwuegbusi, Mark Pett, Lucy Pett, Paul and Christy Russell & Peter Humphreys. Also I thank Nathan, Andy, Jo, Dan, Chris and Amanda for being such wonderful friends.

It would be remiss of me not to mention both my parents for their support. Thanks to my mother for everything but especially for teaching me the joy that can be found in books. Thanks to my father for showing me the things not found in books; honour, honesty, fishing, sport, pride in work and practicality. I hope I have made you proud. Thanks also to my siblings (especially the good looking one!) for always being there and keeping me sane.

Finally, none of this would be possible or even matter without my wonderful wife Davina. Her love and support over the last twenty five years of marriage, through some hard times, some very painful times but also some very good times, has been the inspiration driving me forward. I can never express the debt of gratitude that I owe her or how much she means to me. Davina, THANK YOU, I love you now and always, your loving husband, John.

Abbreviations and Glossary of Terms

Ab - Antibody

aCGH – array Comparative Genomic Hybridisation

ADH - Atypical ductal hyperplasia

Amplification - An increase in the copies of a gene coding for a specific protein without a proportional increase in other genes

AR - Antigen retrieval

BCS - Breast conserving surgery

BM - Basement membrane

CNA - Copy number aberrations

CCL - Columnar cell lesions

CdLOH – Copy deletion loss of heterozygosity

CK – Cytokerratin

CnLOH – Copy neutral loss of heterozygosity

CNV - Copy number variants

Copy Deletion Loss of Heterozygosity - Complete loss of a gene or gene region.

Copy-neutral LOH (CnLOH) - Also known as uniparental disomy (UPD) or gene conversion. An individual receives two copies of a chromosome, or part of a chromosome, from one parent and no copies from the other parent due to errors in meiosis I or meiosis II. This acquired homozygosity may lead to development of cancer if the individual inherited a non-functional allele of a tumor suppressor gene.

DAB - Diaminobenzidene

DCIS - Ductal Carcinoma in Situ

DFS - Disease free survival

Duplication - The repeating of an extra segment of DNA resulting in redundant copies of a part of a gene, a whole gene or sequence of genes

EGFR – Epidermal growth factor receptor

ER – Oestrogen receptor

FEA - Flat epithelial hyperplasia

FNA - Fine needle aspirate

FNAC - Fine needle aspiration cytology

H&E - Haematoxylin and eosin

HUT - Hyperplasia of usual type

IHC - Immunohistochemistry

IMF - Immunofluorescence

LCIS – Lobular carcinoma in situ

LOH –Loss of heterozygosity

MCM - Minichromosome maintenance

MIP – Molecular inversion probe

MRI - Magnetic resonance imaging

NHS - National Health Service

OR - Overall risk

PCR - Polymerase chain reaction

PR – Progesterone receptor

RT - Radiotherapy

Sc gains – Gain of a single copy of a gene.

SNP - Single nucleotide polymorphism

Total Loss - An entire gene sequence coding for a protein is missing from the genome

UEH - Usual epithelial hyperplasia

UPD - Uniparental disomy

Chapter 1. Introduction

5.6 The Normal Breast and Ductal Carcinoma in Situ

The normal human female breast consists of six to ten major duct systems which, along with the lobules, comprise the parenchymal tissue of the breast. Ducts converge via multiple central ducts to the nipple; however each system exhibits little symmetry in shape and size. The ductal system is embedded in stromal connective tissue composed of varying amounts of fat and collagen-rich fibrous tissue. The large ducts and the interspersed smaller terminal ducts terminate in lobular acini. The acini consist of a lumen surrounded by columnar or cuboidal epithelial cells above a single myoepithelial cell layer, which sits upon a basement membrane. The stromal connective tissue around the ducts also contains lymphoid and plasma cells in small quantities. Epithelial cells of the lobular and ductal system are not quiescent, having a high degree of proliferative activity particularly during the secretory phase of the menstrual cycle. Proliferation also increases during pregnancy and lactation and slows down with age and the onset of the menopause.

Ductal carcinoma in situ (DCIS) is defined as a pre-invasive proliferation of cells confined to the breast parenchyma showing no invasion into surrounding tissues, or metastasis (1). The ducts become populated with a clonal expansion of epithelial cells that line, or fill, the breast duct. Cells exhibit the cytological features of malignancy but do not invade across the basement membrane. DCIS is believed to originate in the terminal duct lobular units and not specifically in the larger ducts (1).

1.2 Epidemiology of Ductal Carcinoma in Situ

1.2.1 Incidence of ductal carcinoma in situ (DCIS)

The incidence of DCIS has risen from 1- 5% of lesions detected prior to mammographic screening to approximately 20 - 25 % of all screen detected lesions in the UK (1-3). DCIS may be described as intraductal carcinoma or non-invasive tumour but regardless of terminology it now has the highest increase in frequency of all breast cancers (4). The reason for this seeming rise in incidence is almost entirely down to breast screening mammography, where, in some series, almost 50% of detected lesions are DCIS (5). In some screening populations DCIS is associated with age, where a younger screening population has a higher incidence of the disease than older women (3).

Breast cancer worldwide accounted for 22.9% of all female cancers (1,384,155 cases) and 13.7% (458,503 cases) of all cancer deaths according to the World Health Organisation's figures for 2008. Currently there are no worldwide statistics available on the incidence on DCIS. Countries in the developed world utilising breast screening programs would, by default, register a higher incidence than developing and third world countries, thus making direct comparison difficult. Statistics released by the National Health Screening Breast Screening Program (NHSBSP) for 2010/11 provides data for the incidence of breast cancer and DCIS in the screening population (women aged 47-97 years) for the United Kingdom (6); 2,188,608 women were screened with 17,258 cancers detected. Of these, 3522 were non-invasive or microinvasive carcinoma, accounting for approximately 20.4% of screen detected lesions. In the UK, for all ages, 2,221,975 women underwent breast screening with 14,177 (0.8%) breast cancers detected, with 3613 (25%) of these being classed as non-invasive cancers. There is currently no data available for the type of non-invasive disease, i.e. either DCIS or lobular carcinoma in situ (LCIS), but it is likely that the majority of these will be DCIS, due to its tendency to present with microcalcification

(see below) whereas LCIS is often identified co-incidentally within a biopsy. Cancer Research UK gives figures of 5,765 women and 26 men diagnosed with in situ breast carcinoma in 2010 (7).

In the USA, the incidence of DCIS increased sevenfold from the 1970's to the 1990's. The National Institutes of Health (NIH) Consensus and State of the Science Statement on the Diagnosis and Management of DCIS gave an estimation of approximately half a million women living with DCIS in 2005 (8). The prevalence is higher in Caucasians compared to other ethnicities. A Surveillance, Incidence and End Results (SEER) publication for the USA gives an incidence for in situ cancer of the female breast from 1973 to 2007 for all ages of 21.19 cases per 100,000 (9). This figure rises to 53.42 per 100,000 in the over fifty age group. The incidence of in situ disease in black women and white women is essentially similar in those aged over 50, at 53.41 and 53.35 per 100,000 respectively, but is reportedly higher for white females below 50 years, at 9.03 compared to 7.29 per 100,000. Of note, the NIH advises caution interpreting these data as trends in pathological reporting and definitions have changed over time; as an example, one relevant element may be the absence of data on atypical ductal hyperplasia, which may have been previously reported as benign. However, another review of the SEER data from 1973 to 2000 by Joslyn (10) also identified that women between the ages of 35-44 have a higher incidence of DCIS than older groups. Differences between ethnic groups were also apparent with Asian women more likely to be diagnosed with DCIS followed by black women, white women and finally native Alaskan or American Indians. Thus the NIH data and SEER data are contradictory when comparing the incidence between black and white women but the relative times frames, changes in mammographic breast screening and diagnosis patterns are also possible causes of this discrepancy and highlight the difficulty in analysis of DCIS incidence.

1.2.2 Ductal Carcinoma in Situ and Gender.

Breast cancers are significantly more common in women than men with only 1% of breast cancers being found in males (11). Of this 1% of male lesions it is estimated that approximately 5% are DCIS. The lower incidence can be almost entirely attributed to the increased detection of DCIS in asymptomatic women due to breast screening programs and the absence of screening in the male population.

1.2.3 Risk Factors of Ductal Carcinoma in Situ

Risk factors for DCIS are difficult to quantify due to the absence of data prior to breast screening programmes and the difficulty in defining the asymptomatic incidence in the general population. Risk factors for DCIS are understandably believed to be similar to those known for invasive breast cancer. BRCA mutations are a known risk factor for invasive breast cancer with the cumulative risk by age 70 being 65% for BRCA1 and 45% for BRCA2 mutation carriers (12). Smith et al (12) found that patients carrying the BRCA1 mutation had a lower incidence of DCIS than found in sporadic cases. This may indicate that DCIS has a shorter or 'not observed' phase in BRCA1 mutation carriers. Conversely, Hwang et al (13) found that DCIS was equally prevalent in BRCA mutation carriers as in non-carriers although the onset of DCIS was at an earlier age in the former.

Comparison studies of risk factors for DCIS and invasive breast carcinoma have shown similarities in overall risk (OR). A family history of breast cancer and early menarche showed an increase OR (14). Late childbirth and nullparity however only showed an increased OR in older patients (>50 years) for both DCIS and invasive breast cancer (14). Invasive breast cancer risk is strongly associated with increased age, whereas DCIS is associated comparatively with younger age (14). This would fit with the hypothesis that DCIS is a precursor lesion of invasive disease.

Kerlikowske et al (14) found that increased body mass index (>25kg/m²) resulted in reduced OR in younger females (<45 years) for DCIS with a "trend bordering on

significance". They hypothesise that this is due to composition of breast tissue; women with lower body mass indices tend to have mammographically dense breasts and a two-threefold increased risk of invasive breast cancer compared to those with less dense, fatty tissue found in breasts (as seen in women with a higher body mass index). No increase in risk was found for DCIS in association with alcohol intake, smoking or previous smoking, hormone replacement therapy or oral contraception in individual studies (15-19).

1.2.4 Risk of Ductal Carcinoma in Situ Progressing to Invasive Disease.

The natural history of DCIS remains unclear. When DCIS has been detected subsequent treatment is generally aimed at complete excision of the lesion. Studies focussed upon misdiagnosed DCIS as a benign lesion that could be used to look at the natural history are rare, as are series where the patient has declined surgical excision after definitive diagnosis of DCIS.

The Nurses' Health Study is the largest epidemiological study of women in the USA. It consists of two cohorts funded by the USA's National Institutes of Health, the first started in 1976 by Dr F Speizer (20) and the second in 1989 by Dr W Willett (20), with data on over 120,000 registered nurses. The study provides data on risk indicators for cancer, as well as other diseases. Collins et al (21) reviewed this cohort looking for the incidence of untreated DCIS finding only thirteen cases, thus highlighting the rarity of such incidences. Observations of the natural progression of these untreated DCIS found that ten women developed either a subsequent ipsilateral invasive or an in-situ lesion (2-5 years post original biopsy). The original DCIS lesions included lesions of low, intermediate and high nuclear grade. Although the number of cases studied is small the incidence of 'recurrences' is highly indicative of DCIS presenting a high risk of presenting with a subsequent cancer. In 1978 Betsill et al (22) reviewed intraductal carcinomas treated by biopsy alone and reported that an invasive recurrence occurred in 39% (range 25%-75%) of cases reported in the literature,

however, some recurrences were quoted at ten months and could arguably represent disease that was present at the time of initial DCIS diagnosis.

1.3 Detection of Ductal Carcinoma in Situ

1.2.1 Mammography

The primary screening tool for the detection of DCIS is mammography. Other techniques such as magnetic resonance imaging (MRI), ultrasound and scintimammography are generally regarded as insufficiently sensitive, in the absence of invasive disease (23). Mammograms are radiographic images of the breast that utilise low energy x-rays to identify the contrasting difference between normal breast parenchymal tissue, adipose tissue and breast tumours (1). Radiographic images are usually taken through two aspects, the cranio-caudal and medial-lateral oblique views, to determine the exact location of any lesion. Calcification seen by mammography can be, but is not exclusively predictive of, an indicator of malignant change (see Figure 1).

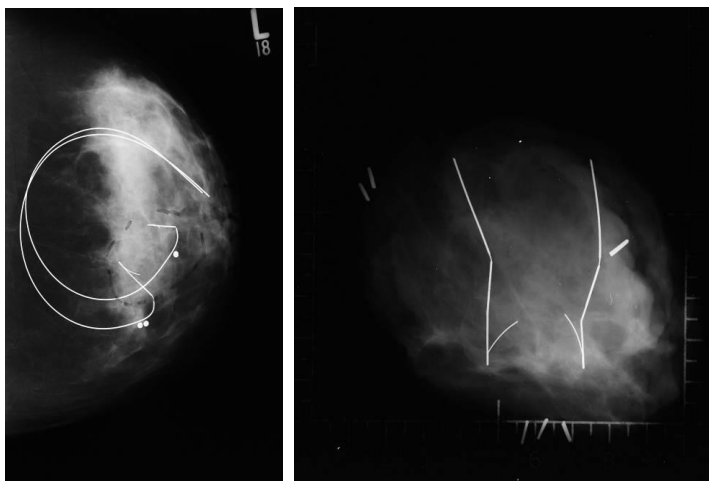


Figure 1: Cranio-caudal and medio-lateral oblique mammographic views of a case with two needle-localisation wires in place to mark an area of abnormality.

The most common mammographic feature of DCIS is microcalcification, seen in 80-90% of cases radiographically (24). Mammographic detection of DCIS thus typically relies upon the identification of calcification deposits that either result from a breakdown product of cells, seen as comedo-type necrosis, or from secretions in the luminal spaces. The clinical presentation of symptomatic cases is variable and may appear as a palpable mass, nipple discharge or as Paget's disease of the nipple (23). Radiographic abnormalities are also often seen in the presence of symptomatic disease with up to 50% of Paget's disease also showing some form of radiological abnormality (23).

Mammographic calcifications (and histological microcalcifications) are not specific for DCIS and occur in many other processes, including normal breast tissue and both benign and malignant lesions, such as involutional changes, duct ectasia and fibrocystic change. It is not possible to accurately grade DCIS on mammography alone and even high grade DCIS lesions can be missed if the area is small.

1.2.2 Ultrasound

Ultrasound is a useful imaging tool for women who present with symptomatic breast disease and can be used to differentiate between a benign and malignant mass in some cases, for example by assessment of the regularity of the margin of the lesion. Its use for identification of DCIS is limited due to a low sensitivity for this disease (1). Nevertheless, ultrasound may be of use in the detection of DCIS where calcification is not present (approximately 16% of cases) which are undetected using mammography. A small study, of 22 cases, has shown that high definition ultrasound can improve upon mammographic screening where calcification is not present (25). Practically, however, it would be difficult to implement ultrasound screening for every patient who attends for breast screening who does not exhibit calcification in the breast on mammography.

1.2.3 *Magnetic Resonance Imaging (MRI)*

Magnetic resonance imaging (MRI) is a highly sensitive method for detection of invasive breast carcinoma. Historically, the sensitivity for benign lesions and for DCIS has been shown to be less reliable than other techniques (1, 2, 26) but more recent studies have indicated that it is more sensitive than mammography (1, 27). MRI does not visualise calcification but uses a contrast agent to detect neovascularisation. Identification of soft tissue changes in DCIS has been reported in 27% of cases using analogue mammography and 60% using digital mammography (27). The cost and time consuming nature of MRI coupled with the claustrophobic nature of the process, however, also inhibits its acceptance as a population screening technique for those not at high risk of developing cancer.

1.3 Histology of Ductal Carcinoma in Situ

1.3.1 *Histological Diagnosis of Ductal Carcinoma in Situ*

Once a breast lesion has been identified and sampled it requires histological diagnosis and classification. Specimens are placed in a fixative or snap frozen before undergoing microtomy followed by haematoxylin and eosin staining (H&E) of cut sections. H&E staining provides morphological and nuclear detail of cellular structures and is a reliable stain for a diagnosis of a particular disease state.

The biological events in the transition from a normal to a malignant state in breast tissues are difficult to assign into predefined groups. There is a spectrum of breast lesions with variable, but low, risk of developing invasive disease. Hyperplastic and atypical changes, such as usual epithelial hyperplasia (UEH), columnar cell lesions (CCLs) including flat epithelial atypia (FEA), and atypical ductal hyperplasia (ADH), represent a disparate group of entities that may exhibit some of the features of, but fall short of, DCIS. These entities are included here to highlight possible diagnostic pitfalls and to demonstrate the morphological and molecular similarities with DCIS.

1.3.2 *Usual Epithelial Hyperplasia*

Usual epithelial hyperplasia (Figure 2) is defined as a proliferation of cells above the basement membrane (BM) that have only mild or no abnormal architectural and/or cytological features. In the normal breast there is a two cell layer above the BM, that of a single layer of myoepithelial cells and a layer of epithelial cells. Where the intraluminal proliferation is confined to four layers or less then the term mild usual epithelial hyperplasia is used. Where five or more cells are present above the BM the term used is moderate epithelial hyperplasia. Where cells have undergone significant proliferation and crossed the ductal spaces enlarging the ducts then the term florid may be used. Clinically there is no increased risk of developing breast cancer where mild epithelial hyperplasia is present and the risk due to moderate or florid usual epithelial hyperplasia is low. Features that distinguish usual epithelial hyperplasia from more serious lesions such as atypical ductal hyperplasia (ADH) or DCIS as set out by the NHSBSP guidelines for 2005 are:

1. The presence of a mixed cell population
2. A syncytial growth pattern
3. Irregular lumen often with a slit-like shape
4. Peripheral lumina
5. Streaming epithelial bridges
6. No abnormal mitoses and few normal ones

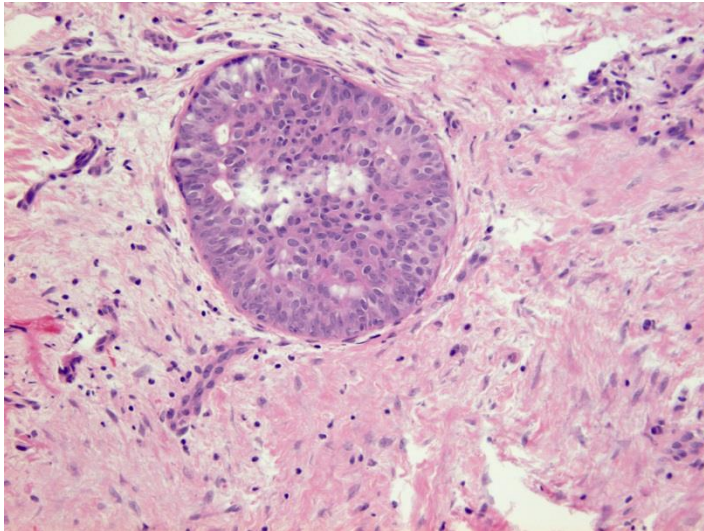


Figure 2: Usual epithelial hyperplasia

1.3.3 Columnar Cell Lesions (CCLs) including Flat Epithelial Atypia (FEA)

Columnar cell lesions (CCLs) are formed from enlarged terminal duct lobular units typically with dilated acini lined by columnar shaped epithelial cells. These lesions are increasingly seen in image guided needle biopsies targeting mammographic microcalcification, which is typically present within secretions in the luminal spaces. In 2003 Schnitt and Vincent–Salomon (28) proposed a classification system consisting of four categories: columnar cell change, columnar cell hyperplasia, columnar cell change with atypia and columnar cell hyperplasia with atypia. The latter two categories were combined to a single entity of flat epithelial atypia (FEA) in 2003 by the World Health Organisation (WHO) Working Group on the Pathology and Genetics of Breast Tumours. Columnar cell change and columnar cell hyperplasia are not considered to increase a risk of breast cancer development. FEA lesions are being detected with increased frequency due to breast screening programs where they may present with mammographic microcalcifications (29). The natural history of these lesions is poorly understood but possible progression to ADH, low grade DCIS and invasive carcinoma cannot be ruled out (29). This necessitates the distinction of FEA (Fig 3) from non-atypical CCLs for future patient management and surveillance.

Molecular evidence indicates that FEA may be the earliest morphological identification of neoplastic change in the breast (30), as FEA lesions have a loss of heterozygosity (LOH) matched by those seen in adjacent concurrent DCIS or invasive cancers. In addition, common chromosomal alterations have been identified on 16q, 11q and 13q, which are not observed in UEH but are frequently seen ADH and DCIS (30, 31).

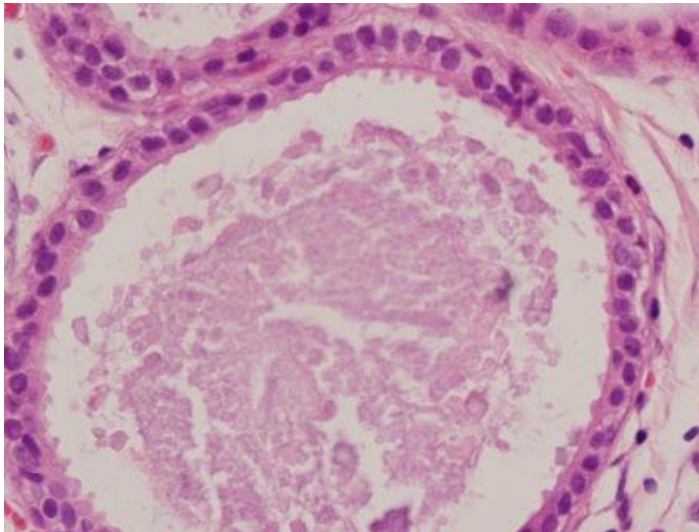


Figure 3: Flat epithelial atypia

1.3.4 Atypical Ductal Hyperplasia (ADH)

The morphology and cytological features of ADH are similar to low grade DCIS but this is a microfocal lesion that is generally confined to a single lobular unit. Specifically, ADH does not involve two complete membrane-bound spaces. The nuclei in ADH are uniform, small, regular and evenly-spaced. As well as the uniformity of cells, ADH forms rigid cellular bars and secondary spaces, which are not seen in FEA (Figure 4). There are presently no immunohistochemical markers available to distinguish ADH from low grade DCIS.

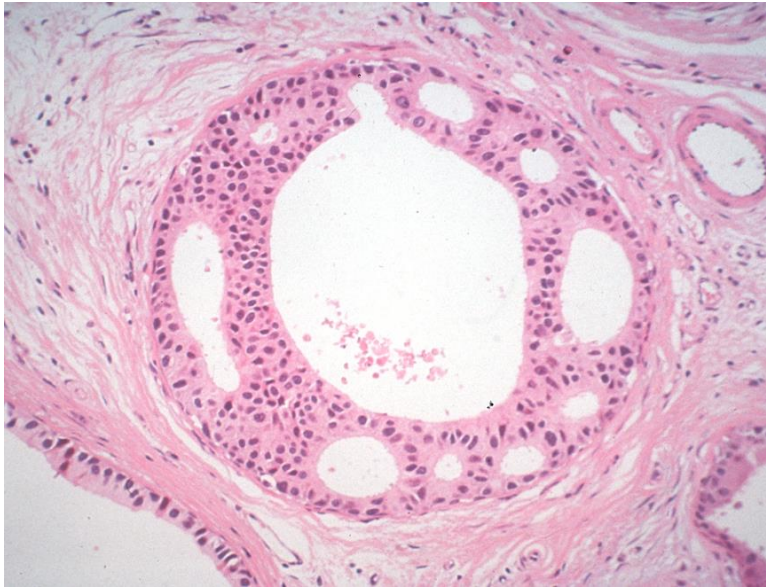


Figure 4: Atypical Ductal Hyperplasia

1.3.5 Lobular Carcinoma In Situ, Atypical Lobular Hyperplasia, Lobular Neoplasia

Atypical lobular hyperplasia and lobular carcinoma in situ have identical cytological features and are distinguished from each other by the extent of involved spaces (1). As such, the term lobular neoplasia was coined to describe both lesions. Cells lack cellular adherence, exhibit eccentric nuclei and often have intracytoplasmic mucinous vacuoles (Figure 5) (1). Lobular carcinoma in situ is considered a separate morphological non-obligate precursor to DCIS, conferring less risk of progression over a longer latent phase.

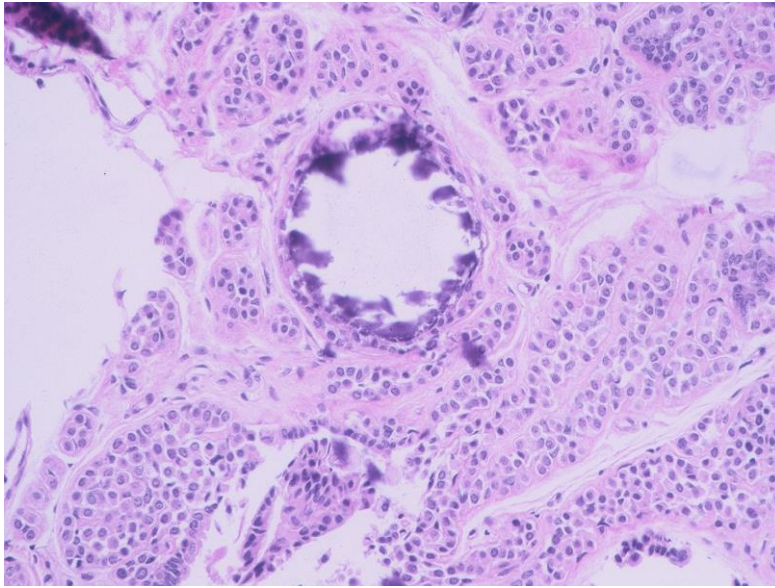


Figure 5. Lobular carcinoma in-situ

1.3.6 *Microinvasive Carcinoma*

Where a DCIS lesion exists and one or more foci of invasion are present but none measure more than 1mm in diameter the classification of microinvasion is given. This is most commonly seen in association with high grade DCIS but is uncommon in DCIS of any grade. Microinvasion may be misdiagnosed when DCIS spreads into lobules, a term known as 'cancerisation' of lobules. Correct classification can often be given by examination of further tissue sections ('levels') and/or immunohistochemistry with myoepithelial and basement marker examination. Cancerisation remains within the breast parenchymal structures and does not extend through the myoepithelial layer and basement membrane (Figure 6).

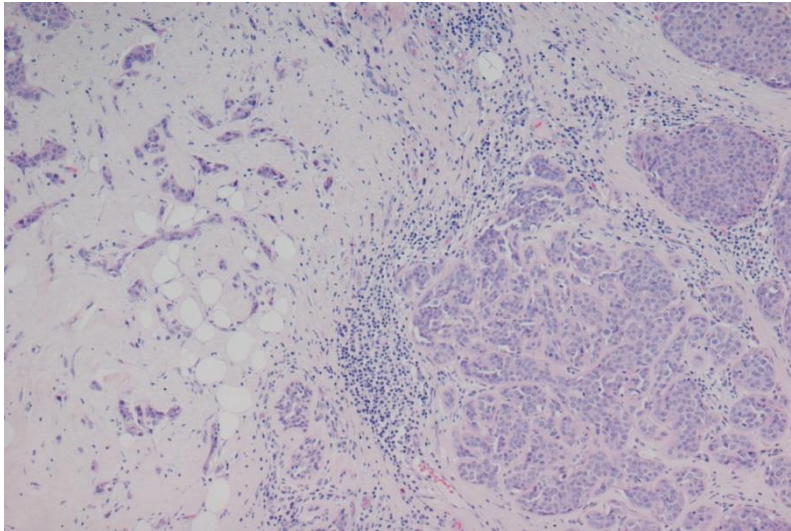


Figure 6: DCIS (top right) with microinvasion (left).

1.3.7 *Histological Variants of Ductal Carcinoma in Situ.*

Histologically DCIS was classified by the architectural pattern of the proliferating epithelial cells. Both the cytology and architecture of DCIS are complex and giving credence to DCIS as a group of heterogeneous lesions. In the United Kingdom the NHS Breast Screening Programme has issued a “Pathology Reporting of Breast Disease Guideline” (32) outlining criteria for the classification and reporting of all breast disease, including DCIS. A brief description and microphotograph of each entity is given below (1, 5, 32).

Comedo pattern DCIS (Figure 7), also historically known as comedo carcinoma, is characterised by sheets of high grade malignant cells with an area of central necrosis typically surrounded by apoptotic cells. The necrotic area often contains microcalcifications identified mammographically. “Comedo” describes the central necrosis, rather than a specific cell type or grade and is therefore not now regarded as a true architectural type of DCIS.

Solid DCIS (Figure 8) is characterised by sheets of malignant cells completely filling the duct space. Cells may be small with few mitoses or, more commonly, large and pleomorphic. Solid DCIS can be confused with lobular carcinoma in situ (LCIS), however the cells of solid DCIS have cell-cell adhesion, seen as distinct cell

boundaries, and may bear microacini with cells showing polarisation around small luminal spaces.

Papillary DCIS/papillary carcinoma in situ (Figure 9) presents architecturally as finger-like projections into the ductal lumen. These are composed of fine fibrovascular fronds surrounded by columnar carcinoma cells. Duct spaces can be expanded/dilated. Myoepithelial cells are not observed between the epithelial cells and the fibrovascular cores, but are present around the external aspect of the duct space in true in situ disease. Where the projections are composed of epithelial cells alone and lack a fibrovascular core the term micropapillary is used. Micropapillary DCIS often also forms complex intraductal bridges and maybe admixed with cribriform patterns.

Cribriform DCIS (Figure 10) has intraepithelial luminal spaces that are classically evenly-distributed and regular in shape. The central area of the lumen is usually present and may or may not contain areas of microcalcification and necrotic material.

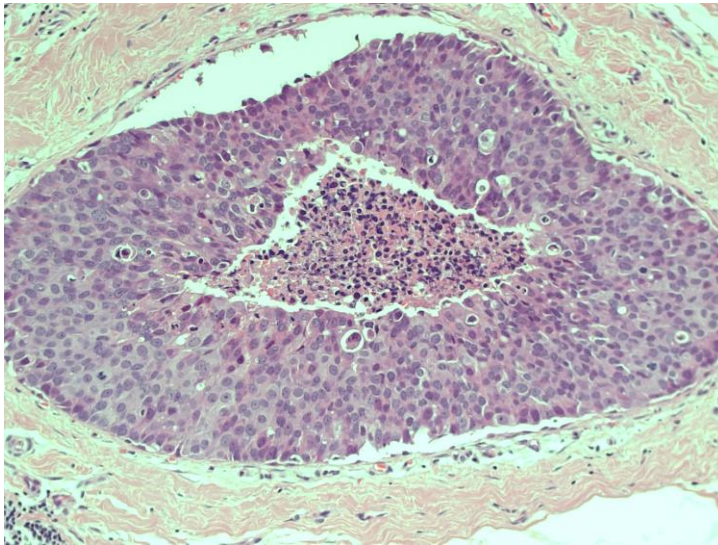


Figure 7: Central comedo necrosis within DCIS

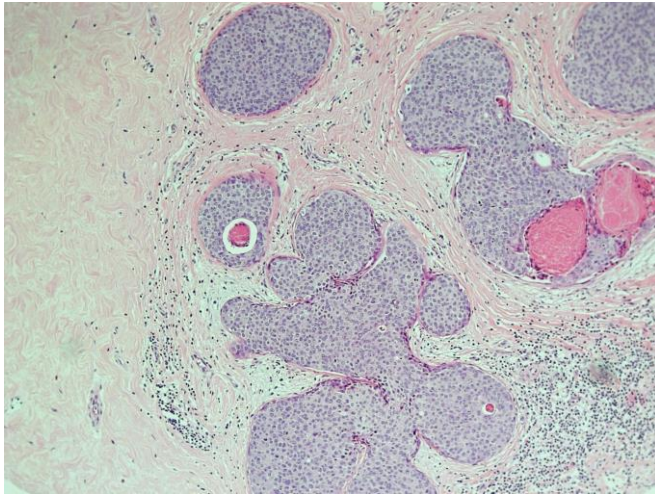


Figure 8: Solid architecture DCIS

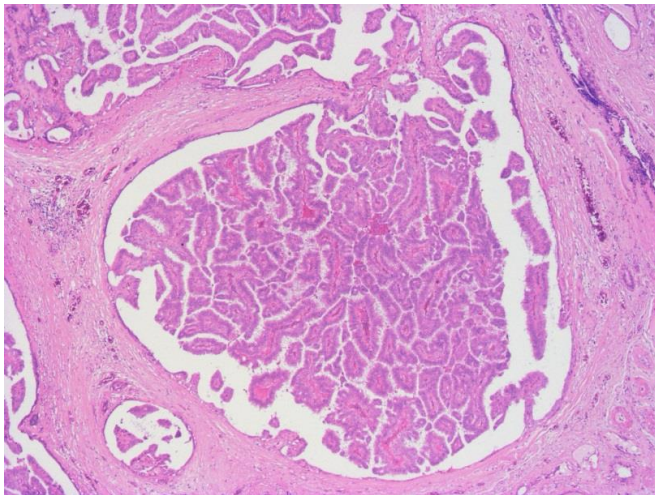


Figure 9: Papillary DCIS/papillary carcinoma in situ

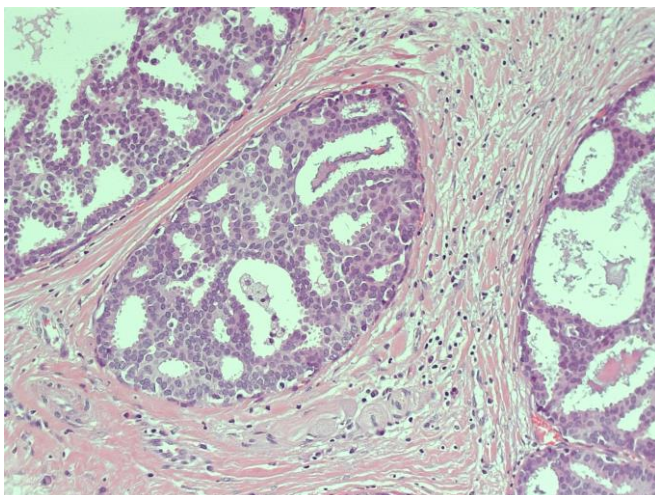


Figure 10: Cribriform DCIS

1.3.8 Rare Histological Forms of Ductal Carcinoma in Situ.

Several rarer morphological variants of DCIS exist, some of which exhibit features that have potential for causing errors in histological diagnosis. The clinical significance of the histological features and genetic mutations present in such lesions remains poorly researched and are not understood.

Clear cell DCIS is formed of cells with distinct margins and clear cytoplasm. Both cribriform and solid structures may be present with central necrosis. Poorly fixed tissues can take on the appearance of clear cell lesions (1).

In apocrine DCIS the cells have the features of apocrine cells with abundant, granular eosinophilic cytoplasm, large nuclei and prominent nucleoli. Severe cytological atypia is often present, as is central necrosis (1).

Neuroendocrine DCIS has polygonal or spindle shaped cells that usually have granular cytoplasm. A pseudo-rosette architecture and papillary areas may be seen. Often positive for oestrogen receptor and neuroendocrine markers this lesion should not be confused with epithelial hyperplasia which has a similar “streaming” architecture (1).

Flat high grade DCIS is possibly a variant of micropapillary DCIS. The features are similar to a variety of columnar cell changes seen in both benign and malignant breast disease which makes identification and classification difficult (1).

Cystic hypersecretory DCIS is also a variant of micropapillary DCIS but produces eosinophilic or mucinous secretions, which expand the ducts to give a cystic appearance. Microcalcifications are common (1).

One of the challenges encountered with classification based on architectural features lies with the presence of more than one of the morphological subtypes coexisting in the same lesion. Descriptive histology terms such as comedo, cribriform, micropapillary, solid, papillary or even “clinging” provide some indication of extent and likely behaviour of disease but are often present in combination, making

categorisation difficult for the pathologist. Quinn et al (33) found that heterogeneity or a mixed architecture was present in up to 62% of cases.

1.4 Classification Systems for Ductal Carcinoma in Situ

1.4.1 *Growth Patterns for Ductal Carcinoma in Situ*

DCIS usually presents in a single duct system and rarely exhibits a multifocal distribution (34). Several studies have been carried out looking at growth patterns in DCIS (35-37). All postulate that where apparently multicentric disease (i.e. more than one cancer separated by normal breast tissue) is present it is due to large lesion size with extended growth along the ductal system. Micropapillary DCIS, however, has been shown to have a greater incidence of multiple foci when compared to other histological subtypes (37).

1.4.2 *High Grade, Intermediate Grade and Low Grade Ductal Carcinoma in Situ*

There are several grading systems (38-43) classifying DCIS largely utilising cytonuclear features, into high grade, intermediate grade or low grade categories. It should be noted that whichever grading system is used that there is a degree of interobserver variability between pathologists, particularly with regard to intermediate grade DCIS. The current UK grading system for is based upon guidelines set by the NHSBSP (6).

High Grade DCIS is characterised by large pleomorphic cells with variation in size shape and polarity. Coarse chromatin and prominent nucleoli are abundant and mitotic cells are frequently seen. Nuclear outlines have irregular contours and may appear crenelated with the enlarged nuclei being greater than three times the size of erythrocytes. High grade DCIS is usually associated with a central necrosis and

calcification deposits within this area. This calcified necrotic debris is identified on mammographic screening.

Low grade DCIS comprises more evenly-spaced uniform cells that are generally smaller than other forms of the disease. The cells have regular, round nuclei located centrally and are normally one to two times the size of erythrocytes. Cells are arranged in well-ordered patterns showing polarity. Mitotic figures are sparse. A solid sheet architecture is rare but cribriform and micropapillary patterns are common, often with both present in the same lesion. When both are seen the cribriform pattern tends to dominate.

Intermediate grade DCIS cells exhibit moderate pleomorphism and lack the uniformity of a low grade lesion. Nuclear size is greater than that seen in low grade DCIS being two to three times the size of erythrocytes and occasional nucleoli are present. Lesions can have a cribriform, solid or micropapillary pattern with some polarisation of cells.

Although rare, sometimes a DCIS lesion will exhibit a range of nuclear features; when this occurs the classification is according to the highest nuclear grade present.

1.4.3 The Van Nuys Grading System for DCIS and the Van Nuys Prognostic Index (VNPI)

The Van Nuys grading system (38) applies a combination of cytonuclear features (as above), combined with the presence of comedo-type necrosis to classify DCIS. High grade lesions are classified according to the cytonuclear grade, regardless of the presence of necrosis, whilst non-high grade lesions are classified as non-high grade with necrosis or non-high grade without necrosis. This classification divides DCIS into three groups with differing risk of recurrence.

The grading system was further enhanced into a prognostic index for DCIS by adding tumour size and margin width (distance of tumour cells from excision margin) to the pathological classification. This classification was named the Van Nuy's Prognostic

Index (VNPI) (39). The VNPI takes into account these three parameters associated with local tumour recurrence in patients with DCIS. Tumour size, margin width and pathological classification are all assigned a score of 1-3; 1 relating to best prognosis. A sum of the scores, of 3-9, was determined and a cut-off point relating to treatment failure, i.e. recurrence, was calculated. The original classification study was carried out on 333 patients and determined that a VNPI score of 3 or 4 showed no additional benefit for disease free survival from radiation treatment compared to scores of 6 or 7 who did receive benefit. Those patients with scores of 8 and 9 had a high recurrence rate regardless of treatment. The VNPI was updated by Silverstein et al as the University of California/ VNPI in 2003 with the addition of age as a factor in determining treatment (44).

1.4.4 The Tavassoli Grading/Classification System

The Tavassoli system (40) splits high grade, intermediate and low grade DCIS into ductal intraepithelial neoplasia groups (DIN). The designation removes 'carcinoma' from the terminology, similar to neoplasia grading in cervical screening. DIN3 represents high grade DCIS with, or without, marked intraluminal necrosis and severely atypical nuclei. DIN2 is characterised by cribriform or micropapillary type lesions with mild to moderate nuclear atypia, with or without necrosis. The classification of DIN1 requires the absence of necrosis and mild atypia with uniform nuclei and includes hyperplasia of usual type (HUT) and atypical hyperplasia (see above).

1.4.5 R Holland and the European Pathologists Working Group Classification System

This grading system, first proposed by Holland et al (41), uses nuclear pleomorphism combined with cell polarity to define high grade, intermediate grade and low grade DCIS. Polarisation of cells is the orientation in relationship to luminal spaces, where

loss of polarisation indicates a higher risk of local recurrence. The European Breast Screening Working Group adopted a similar system (42), relying on nuclear grade but ignoring polarisation on the grounds of difficult reproducibility amongst pathologists. This is the system currently recommended in the NHSBSP guidelines (32).

1.4.6 Summary of Classification of DCIS

There are at least ten grading systems described for reporting of DCIS. All have their merits and disadvantages. Shoker and Sloane (45) reviewed the available classifications in 1999 and determined that nuclear grade was the best predictor of risk of recurrence. Pinder et al (46) reviewed over 1600 cases of DCIS for the UK Coordinating Committee on Cancer Research (UKCCCR) trial. A full pathological assessment of all cases was undertaken and compared the Van Nuy's, Nottingham and Holland classifications. Their results found that all grading systems of DCIS showed a significant association with recurrence of ipsilateral DCIS or invasive disease. However, they identified an additional subgroup, which had a particularly poor recurrence rate. This group was defined by over 50% of ducts showing central comedo type necrosis, a high cytonuclear grade and over 50% of ducts showing a solid architecture pattern. Pinder et al proposed a new three group system to define DCIS. The categories are: 1) low and intermediate cytonuclear grade DCIS, (2) high cytonuclear grade DCIS with non-solid architecture and < 50% of ducts with necrosis (3) high cytonuclear grade DCIS with a predominantly solid architecture and >50% of ducts bearing central comedo necrosis.

All of the classification systems rely upon traditional pathological evaluation of H&E stained tissue sections. Whilst they all have been shown to have a predictive element and be related to risk of disease recurrence, none are currently definitive in the assessment of a possible relapse; thus even patient's with low grade (low risk) DCIS may develop recurrent disease as in situ or invasive breast cancer. This may be due

in some part to the subjective nature of pathology reporting and interpretation of morphological definitions.

1.5 Treatment of Ductal Carcinoma in Situ

1.5.1 *Surgical Procedures and Options for Ductal Carcinoma in Situ*

Treatment options for DCIS are similar to those for invasive carcinoma being mastectomy, or breast conserving surgery (BCS) with or without radiotherapy (RT) (47). The American College of Radiology (ACR) issued appropriateness criteria in 2011 which reflect both the heterogeneity of DCIS and the patient-driven choice (47). The recommendations state that BCS with whole breast radiotherapy for localised DCIS is acceptable, as opposed to mastectomy. Older patients with fully excised low grade DCIS may undergo observation alone. In cases where microinvasion is present with DCIS (mDCIS) then a sentinel lymph node biopsy is advised.

1.5.2 *Radiotherapy for Ductal Carcinoma in Situ*

Radiotherapeutic treatment for DCIS is primarily recommended after breast conserving surgery and seldom performed after a mastectomy for DCIS. There remains some controversy as to the clinical need and effectiveness compared to breast conserving surgery alone. In 2013 Donker et al (48) reviewed the treatment of DCIS with and without radiotherapy (RT) after local excision of lesions. They studied survival rates over a period of fifteen years for both groups from a European Organisation for Research and Treatment of Cancer (EORTC) randomized phase III trial (number EORTC10853). They found that 30% of women who underwent local excision for DCIS without radiotherapy had a local recurrence (LR) with 50% of these being invasive. In the group in receipt of both local excision and RT, only 17% had a recurrence, with 56% of these having an invasive recurrence. In both groups the risk of recurrence in the first five years was greater than ten or fifteen years. From these data radiotherapy is shown to decrease the risk of local recurrence. However, a

difference in local recurrence between the two groups was not translated into any change in breast cancer specific survival. All subgroups of DCIS had a reduced recurrence rate with RT with clinging/micropapillary DCIS showing the least additional benefit (no RT = 12.5% recurrence, with RT = 5%).

Other clinical trials have similarly demonstrated a reduction in local recurrence with the use of RT (49-52). In a Swedish study, Holmberg et al (49) found that RT substantially reduced the risk of recurrence but was more beneficial in older women compared to younger women. Confounding factors in the treatment of DCIS with RT such as DCIS grade, age, RT dose, length of dosage time and the use of boost (a higher final dose aimed at the tumour bed) and the effect of adjuvant intervention (such as hormone therapy) make definitive conclusions difficult to make over such long survival periods. The Loris trial (53) is an ongoing a phase III trial addressing possible overtreatment of low grade DCIS by using active monitoring as opposed surgical treatment for screen detected or asymptomatic low grade DCIS.

Cuzick et al (50) looked at the effect of Tamoxifen and RT in DCIS patients. 1701 women were randomly assigned to treatment groups: RT and Tamoxifen, RT alone, Tamoxifen alone, or to no adjuvant treatment. Their findings also showed a benefit in reducing either DCIS or invasive ipsilateral recurrence with RT. They also found that RT had no benefit for overall survival (OR). They reported that Tamoxifen had an increased benefit in prevention of development of cancer in the contralateral breast.

1.5.3 Hormone Therapy for Ductal Carcinoma in Situ

DCIS may be treated with anti-oestrogens such as Tamoxifen or aromatase inhibitors (AI) if a lesion is positive for oestrogen receptors. Staley et al (54) reviewed the Cochrane Central Register of Controlled Trials (CENTRAL, *The Cochrane Library*), the Cochrane Breast Cancer Group's Specialised Registry, and the World Health Organization's International Clinical Trials Registry Platform (WHO ICTRP) in 2011 looking at randomised clinical trials using Tamoxifen with or without radiotherapy.

They conclude that, in a similar fashion to RT, Tamoxifen reduces the recurrence rate but has no effect on OS. The NASBP-24 trial (55) found that Tamoxifen reduced recurrence in ER positive DCIS but not in ER negative DCIS.

1.5.4 Chemotherapy for Ductal Carcinoma in Situ

Chemotherapy as a treatment is not usually given or recommended for DCIS. As pure DCIS is a non-invasive lesion, which has not metastasized, systemic treatment is regarded as having little or no benefit for patients.

1.6 Techniques for the Diagnosis and Investigation of Prognosis in DCIS and Invasive Breast Cancer

1.6.1 Sampling Techniques.

Once a suspected lesion is identified, either through screening or symptomatic presentation, a confirmed histological diagnosis is required. In order to achieve this there are several possible sampling techniques the clinician may choose. These are discussed below with relation to suitability for a diagnosis of DCIS.

Fine needle aspirate (FNA) or fine needle aspiration cytology (FNAC) is a percutaneous process of using a small needle attached to a syringe to extract cells. Cells are directly drawn from the site of interest and smeared onto glass slides where the cells are stained and interpreted. This procedure is suitable for the identification of cancerous cells but has severe limitations in its ability to distinguish in-situ lesions from invasive carcinoma, due to the lack of relevant architecture/morphology. The NHS Breast Screening Programme specifically does not recommend FNAC for the diagnosis of microcalcifications.

A core biopsy is the removal of a solid core of tissue by the use of a hollow needle. The needles can be of various sizes or gauges. The choice of biopsy size for palpable lesions is often a matter of clinical preference, although there are studies that indicate that a 14 gauge is superior to either a 16 or 18 gauge needle (56, 57). The use of

ultrasound guidance and/or a stereotactic needle biopsy (where 3 dimensional coordinates are used to locate the area to be sampled on x-ray) are required for non-palpable lesions. There is evidence that ultrasound guidance reduces the false negative rate in identifying breast carcinoma using core biopsies (58, 59). Core needle biopsies are suitable for the identification of DCIS, however some caution is advised when giving a definitive diagnosis based on core biopsy alone (60); studies of the underestimation of invasive disease where a core needle biopsy shows DCIS range from 8% to 41% (58, 61). Chan et al found that lesion size and the number of cores taken were predictive factors in cases where a diagnosis of DCIS was made but there was adjacent invasive disease present compared to core sampling alone (62).

Vacuum-assisted biopsies utilise a larger gauge needle and have the advantage that multiple tissue samples can be obtained from one insertion. Vacuum-assisted biopsies are normally performed under image guidance, in the form of x-rays or ultrasound, to assist the sampling of breast abnormalities. Studies comparing vacuum-assisted biopsies against non-vacuum stereotactic biopsies have shown that the need for repeat biopsies is reduced, although not eliminated (63, 64).

In some instance definitive diagnosis with FNAC or core biopsy pre-surgery may not be possible. If there is a high suspicion of malignancy then an open biopsy may be performed by surgical excision.

1.6.2 Fixation and Processing of Biopsies

Once a tissue sample has been removed, it requires preparation for light microscopy. The standard practise in most UK laboratories is to fix tissues in a formaldehyde fixative, or one of its derivatives, and process to paraffin wax blocks. Fixation stops putrefaction and autolysis whilst processing to paraffin wax has a twofold aim. This is the removal of water, which assists fixation, and the hardening of tissue so that it can be suspended in a suitable medium (paraffin wax). This enables slicing thin enough

for light to pass through thus enabling light microscopy. Processing is achieved by passing tissues through a variety of graded alcohols and organic solvents followed by impregnation with molten wax. Thin sections of wax embedded tissue are cut, usually 3-4 μ m thick, and stained with haematoxylin and eosin stain (H&E) to demonstrate the nuclear and cytological components of cells. A diagnosis is generally made on the H&E stains, although a panel of immunohistochemistry (IHC) antibodies may also be used as ancillary tests to aid in a definitive diagnosis. In invasive breast cancer additional IHC is then performed to establish the expression of hormonal markers such as oestrogen and progesterone receptors and growth factors such as HER2 in order to determine subsequent optimum clinical treatment.

Immunohistochemistry (IHC) is a technique for identifying tissue components (antigens) by means of an antibody-antigen reaction. Antibodies specific to a known epitope, or epitopes, can be produced by injecting a host species (usually a mouse or rabbit) with the antigen. The host then produces an immune response in the form of antibodies. These antibodies (known as the primary antibody) can be harvested and applied to histological sections where they will bind to their target antigen if it is present. In fluorescence microscopy a fluorophore is usually bound to the primary antibody to give a fluorescent image. In chromogenic or light microscopy the primary antibody signal is usually amplified by the application of secondary antibodies. These secondary antibodies are links binding to the primary antibody and an enzyme, either with the use of a polymer chain or an avidin biotin complex. The substrate for the enzyme forms an insoluble coloured end product when applied and can be viewed using light microscopy. The most common enzyme substrate reaction used in IHC is currently horseradish peroxidase (HRP) combined with a diaminobenzidine (DAB) substrate, which produces a brown coloured end product. The use of IHC to identify proteins or antigens in tissues has significant advantages over other protein assays, as the target protein can be identified in its morphological setting. This enables identification of proteins in tumours or normal tissues.

IHC as a technique does however have some inherent problems. In the UK, the majority of tissues for examination are fixed in formaldehyde or one of its derivatives. Whilst an excellent fixative, formalin produces cross links in the form of hydroxymethyl and methylene bridges, which may effectively mask the antigen of interest (65), and prevents primary antibody binding. In order to overcome this, antigen retrieval (AR) techniques utilising heat and salt solutions have been incorporated in many IHC protocols. Unfortunately there is little standardisation of the use of fixatives with regard to type and length of fixation, resulting in the need for optimisation of individual antibodies with known positive controls in each different laboratory. The lack of standardisation in laboratories may affect interpretation of IHC markers and is one of the contributory factors that make IHC scoring somewhat subjective in nature. However, despite these limitations, IHC has proven to be a valuable tool in both diagnosis and in prediction of response to therapy in breast cancer diagnosis.

1.6.3 Immunohistochemistry in Assessment and Prognosis of Ductal Carcinoma in Situ

Immunohistochemistry is commonly used to assess DCIS. The demonstration of an intact basement membrane (BM) or myoepithelial cell layer is integral to a diagnosis of DCIS as opposed to an invasive lesion. Antibodies raised against the BM include collagen IV and laminin whilst myoepithelial markers such as smooth muscle myosin heavy chain (SMMHC), p63 and smooth muscle actin (SMA) are all used in diagnostic laboratories to confirm that the lesion is in-situ. Other IHC markers may be used to help in the differential diagnosis of a lesion, for example E-cadherin is typically positive in DCIS but absent or reduced in LCIS, and ER and basal cytokeratins, such as cytokeratin 5 and cytokeratin 14, typically show homogenous staining patterns in DCIS and can be used to aid in a differential diagnosis between UEH and DCIS (1). There are currently no IHC markers to distinguish ADH from DCIS.

The demonstration of hormonal markers, HER2 or other IHC markers are not routinely used to classify DCIS or predict the possibility of invasion. ER expression in DCIS remains difficult to assess mainly due to the lack of clinical data defining positive and negative thresholds. The various scoring techniques employed e.g. Histoscore, Allred and percentage score also prevent systematic comparison. Expression of ER is lower in high grade DCIS than low grade DCIS and the percentage of DCIS ER positive cases is approximately 70%, which is similar to that found in invasive series (1). The expression of HER2 in DCIS has been shown to be higher in high grade DCIS than in grade 3 invasive tumours, with expression in 50-60% cases compared to approximately 20%, cases respectively. For example, Park et al (66) found that pure DCIS expressed HER2 at a greater frequency than invasive breast disease. Borgquist et al used tissue microarrays to determine Her2 expression in primary DCIS. In their cohort of 458 patients over 30% had Her2 positive DCIS (n=132). They found the risk of developing an invasive breast cancer recurrence was “statistically significantly lower subsequent to a HER2 positive DCIS compared to a HER2 negative DCIS(67)”,

1.6.4 Tissue Microarrays (TMAs)

Tissue microarrays (TMAs) were first described in their current format by Kononen as a tool for gene expression and copy number analyses (68). Subsequently, the technique was recognised, and has been widely used, as an efficient method for assessing biological markers on large numbers of cases economically. Biomarker validation using TMAs with immunohistochemistry (IHC) and in situ hybridisation (ISH) is valuable for evaluation of putative diagnostic and prognostic markers and also drug targets when applied to large series and especially to randomised clinical trial samples. Tissue microarrays (TMAs) have become an established research tool over the last ten years enabling rapid assessment of significant numbers of test tissues using tinctorial staining, IHC or ISH. The premise of the procedure is to sample pre-existing blocks of histologically processed tissue for further evaluation by

removing individual cores from each test case and amalgamating these into one block from which multiple tissue sections can be taken for analysis. The procedure allows simultaneous testing of multiple cores, giving reduced costs and improved standardisation of experiments. The vast majority of TMAs are constructed from formalin fixed paraffin wax embedded tissues using 0.6mm, 1.0mm or 1.5mm cores taken from a donor block using a hollow tube (stylet) to biopsy the original tissue which are then are placed into a recipient wax block (Figure 11) into which up to a hundred or more cores can be placed. Once the blocks have been constructed the cores are annealed to the recipient block by a process of gentle heating and cooling. Once this has been achieved, sections are cut in the normal manner on a microtome and subsequent slides can undergo IHC or ISH testing.

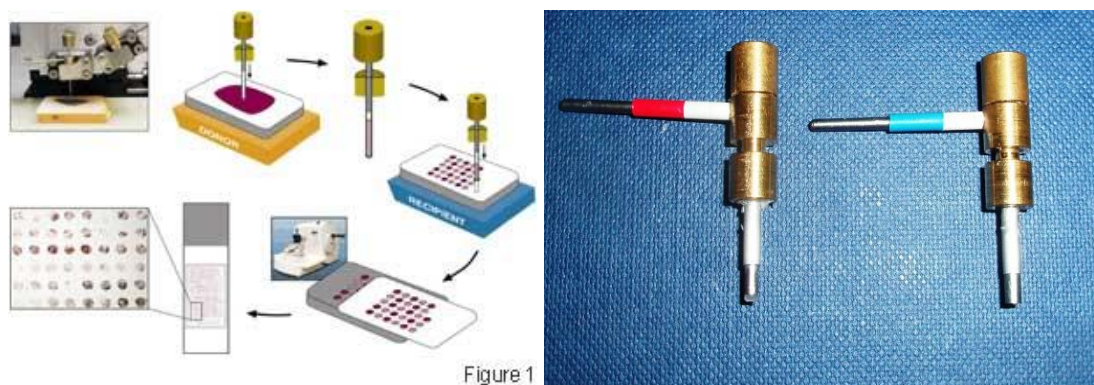


Figure 11: TMA manufacture and stylets

(TMA image courtesy of Yale University, Department of Pathology, Connecticut USA.)

1.6.5 Genetic Aberrations & Molecular Inversion Probe Arrays for Genomic Profiling

Profiling of the human genome is increasing our understanding of the disease process and is seen as a potential tool to identify new biomarker targets for interventional therapy. However, the human genome contains more than ten million nucleotide

positions that have >1% or “common” variations between individuals. Identification of these differences and subsequent comparison of the genomic changes found in progression from normal to a tumour state in breast and other, cancers requires reliable and robust sequencing methodologies. The field of genomics has made significant advances since the publication of the human genome sequence simultaneously by Lander et al (69) and Ventner et al (70).

An individual has two copies of the genome, one from each parent. Where a single base differs in the genome between these two inherited copies a single nucleotide polymorphism (SNP) is present giving rise to heterozygosity in the individual's genome. Chromosomal arrangement in humans is paired, allowing for heterozygosity. However, loss of a genomic region from one parent can occur giving only one region, which by default cannot be heterozygous at a SNP location; this is termed as loss of heterozygosity (LOH). LOH of a tumour suppressor gene results in only one copy of the suppressing gene. A point mutation (replacement of a single base) can lead to inactivation of the tumour suppressing gene given rise to potential tumourigenesis. However, LOH is not restricted to a loss of a genomic region; copy neutral LOH can occur where no overall net loss of chromosomes or copy number is apparent. Also known as unipartietal disomy (UPD) (71), errors during meiosis may result in two copies, or partial copies, of a genomic region which are inherited from one parent and not the other parent. This can cause inherited homozygosity and, if a non-functional tumour suppressor gene is present, may also result in tumorigenesis. It is estimated that up to 50-75% of human cancers demonstrate LOH (71, 72).

In addition to LOH, and copy neutral LOH, deletions of chromosomal regions and amplifications can also lead to cancer development. Copy number variations (CNVs) are abnormal variants in one or more regions of the genome. They usually correspond to large regions ranging from one kilobase to several megabases and can be either deletions or amplifications. They differ from SNPs in that the copy number variation can be repeated along the length of DNA, e.g. a section that has repeat elements with

GGATCC may be replaced with GGGATCC (duplication) or GGATC (deletion). CNVs have been estimated to account for approximately 13% of human genomic DNA (73). Various platforms have been developed to improve the speed, accuracy and quality of genomic profiling (74). The goal of personalised whole genome association studies requires the ability to identify aberrant genomically regions from very little tissue or DNA with minimal cross reaction; as may seen with PCR amplification where carryover of PCR products to subsequent amplification reactions can occur, in a cheap and reproducible system (75). Molecular Inversion Probe (MIP) arrays are a technology for analysis of SNPs, copy number alterations, somatic mutations and loss of heterozygosity (LOH) (including copy-neutral LOH). The probes consist of primers of approximately 20 nucleotides separated by a linker region (usually an endoribonuclease) and sequences of single stranded DNA that are complimentary to regions adjacent to a SNP in the target DNA sequence of a sample. The technique works on the principle of a SNP-genotype reaction being placed directly upon the genomic DNA target prior to any amplification and then subsequently using a PCR step to only increase the amplification of detected target molecules. This alleviates problems of multiplex PCR where undesired reactions such as “primer dimers” are caused by unwanted binding of complimentary bases instead of the intended DNA amplicons. Each MIP has two sequences that are complementary to the target DNA at the 3 primed (3') and 5 primed (5') ends of the DNA strand (Figure 12). Each sequence hybridizes to the target, becoming “inverted” and forming a circle with the 3' and 5' ends adjacent to each other with the restriction endoribonuclease forming a free hanging loop.



Figure 12: MIP Array Probe:

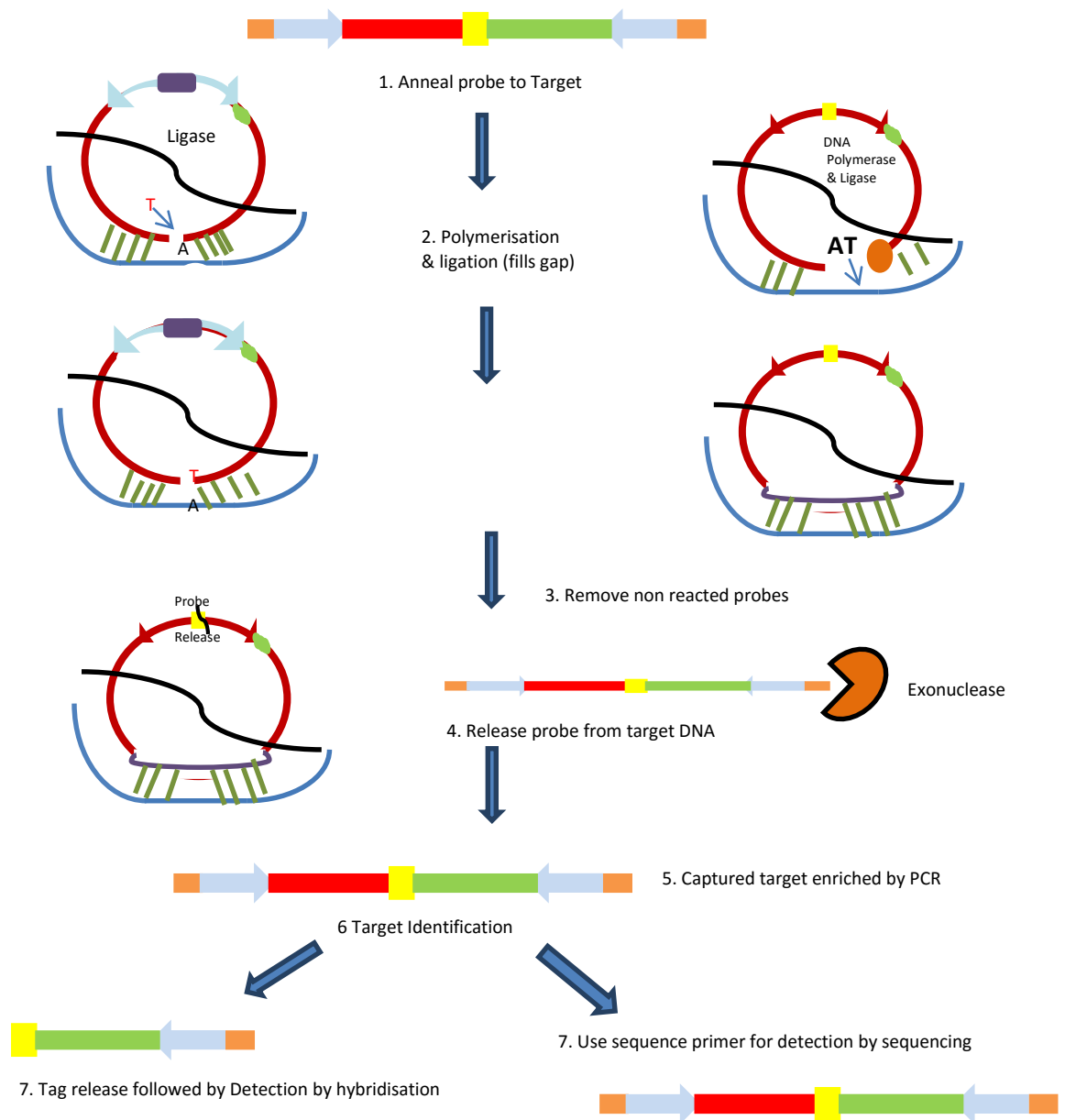


Figure 13: MIP array process for genomic profiling

(Adapted from http://en.wikipedia.org/wiki/File:MIP_technique_timothy.png#filelinks

copyright permission licensed under the GFDL see: Appendix 1)

MIP array technology has been further developed by the commercial company, Affymetrix (California, USA), to be used in archival formalin fixed paraffin wax embedded material (FFPE). The company cites the advantages of MIP array as the ability to identify genomic regions of interest in partially degraded samples such as FFPE tissue allowing as little as 40bp per genomic region of interest and being able

to identify difference in copy number states in as little as 60bp with a high discrimination between “signal” and “noise” ratio in degraded samples. One additional advantage of MIP array technology is the multiplexing ability where thousands of loci can be simultaneously analysed; the Affymetrix system offers over 300,000 markers.

1.7 Molecular Portraits of Breast Cancer

1.7.1 *Genetic Subgroups of Invasive Breast Cancer*

The morphological and molecular heterogeneity of breast tumours makes the classification of lesions challenging. Perou (76) and Sorlie (77) have shown that invasive breast cancers can be classified into distinct genetic subgroups that correlate with overall survival (OS) and disease free survival (DFS) and have potentially important clinical implications. Their findings have been corroborated in other studies (76, 78, 79). The sub-groups of invasive breast carcinoma, designated as luminal A, luminal B, ERBB2-positive, basal-like and “normal” demonstrate that the phenotypic variety seen histologically is, unsurprisingly, present at the molecular level. The molecular phenotypes correlate well with certain immunohistochemical profiles. Approaches have confirmed the existence of some subclasses, such as the basal-like groups of tumours which had been previously described in some IHC series (78, 80-83) but which, until genomic expression profiling, had not been truly regarded as a distinct subgroup of invasive breast cancer.

Perou’s experiments examined breast cancer profiles using known gene clusters for 65 tissue tumour samples along with cultured cell lines of known genomic sequences to ascertain a molecular signature of the breast cancers studied. In developing a system for classifying tumours based upon the gene expression, a second subset of genes (known as the intrinsic subset) to further classify tumours was examined/proposed. Using the rationale that any samples taken from the same tumour should demonstrate a similar profile and that separate tumours should vary in

their expression, twenty paired samples were analysed and gave rise to identification of 496 genes with significant variation between tumours compared to the paired samples. As Perou points out there, is a great multidimensional variation in gene expression in that the patterns are mainly independent. However, there are distinct molecular portraits that relate to both similarities and differences of tumours. These include a proliferation cluster as the largest group, which correlated with the histological mitotic index and the immunohistochemical markers of proliferation, Ki67 and PCNA. As the hierarchical cluster is further broken down, two dendrogram branches reveal oestrogen receptor (ER) positive and ER negative tumours separated by clusters of genes expressed by luminal cells of the breast. The ER positive, luminal grouping was also correlated with immunohistochemical expression of cytokeratins 8 and 18. Further evaluation by Sorlie et al, elucidated that these luminal groups could be further subdivided into groups luminal A, B and "normal" like. Six tumour samples from the original Perou experiments clustered in a group associated with genes characterised in basal or myoepithelial cells of the breast. This group had a protein expression of cytokeratins 5/6 and 17. The basal-like subgroup has particular clinical implications as the vast majority of these tumours exhibit negative IHC staining profiles for ER, progesterone receptor (PR) and HER2. These tumours are unlikely, therefore, to respond to Tamoxifen or aromatase inhibitors as they lack expression of hormone receptors. Similarly, the lack of HER2 expression means that tailored anti-HER2 therapy, such as with trastuzumab, is also likely to be of little value. The lack of expression of these three markers which is frequently seen in these basal-like lesions has resulted in the 'triple negative' term being erroneously used synonymously to such "basal-like" tumours. Rakha (84) reviewed the expression profiles of basal-like lesions and addressed some of the confusion surrounding the definition of basal-like tumours. Rakha points out that there is a large subset of the basal-like group that is triple negative in immunophenotype but that the absence of these three markers does not automatically indicate a basal-like

phenotype. Conversely, tumours within other molecular sub-groups may also exhibit a triple negative profile. It should also be noted that the expression of one single “basal” IHC marker does not necessarily detect a basal tumour. Rakha advocates the use of a panel (CK5/6, CK14, CK17, EGFR), with positive expression of these, as well as ER and HER2 negativity required for the identification of a basal-like tumour. The majority of invasive breast cancers arising in BRCA1 carriers also fall into the basal group, but this also is not exclusive. The definition of basal-like tumours cannot therefore be defined by either a triple negative phenotype or a BRCA1 carrier group, once again demonstrating the heterogeneity of invasive breast tumours. The basal-like immunophenotype has also been identified in forms of DCIS (85-87).

The next subdivision postulated by Perou et al is the HER2 expressing group. This cluster contained not only the Erb-B2 locus, but also several other genes located in the same region on chromosome 17. This subset also showed lower levels of ER positivity.

The “normal like” subgroup was identified and named based on similarities in gene expression patterns to those of normal epithelial cells, adipose tissue and other non-epithelial cell type. Norum describes these tumours in 2014 as “clearly carcinomas that have gene expression features common to both basal-like and luminal subtypes, showing no expression of proliferation-associated genes” There is an acknowledgement that these samples may have a low tumour cell percentage within the samples(88).

The molecular profiling of invasive breast tumours, and subsequent subdivision of tumours into groups, may have wide-ranging clinical implications. Currently, the pathological assessment of tumours and patient’s prognosis and subsequent clinical management is based upon descriptive histological factors such as lymph node stage, pathological measurement of carcinoma size and histological grade, the latter incorporating a proliferative assessment, and morphological (architectural and cytological) subtypes. These factors are routine coupled with the IHC expression of

several markers, notably ER, PR and HER2, which are used as predictive markers and on which likelihood of response to hormone therapy and trastuzumab is based. Other antibodies can help with the identification of morphological tumour type but have a limited prognostic value; for example, it is well recognised that invasive lobular carcinoma typically shows a loss of E-cadherin immunohistochemically. In addition, many novel markers, although demonstrating some potential as prognostic markers, for example c-kit (86, 87), Aurora A (89, 90) and BCL2 (91-93) are not widely accepted as independent predictive markers of significance compared to existing pathological criteria.

The molecular expression profiles demonstrate that morphologically similar tumours may be subdivided into groups which appear to have different genetic origins, leading to questions regarding current histological classification, particularly the “histological special types” (94, 95). Such findings also raise questions as to the validity of current practice. However, it should be noted that studies of genetic molecular profiling of invasive breast tumours have been carried out on relatively small numbers. Validation of these classification criteria using microarray based systems on thousands of samples is cost prohibitive and any novel genetic taxonomy would require detailed correlation with existing clinical practice.

1.7.2 IHC as a Surrogate to Genomic Profiling.

The use of tissue microarrays and IHC can facilitate the examination of the robustness of the genetic subdivisions. Several studies have focussed upon this method, with promising results (82, 86, 96). Abd El-Rehim (96) et al analysed over 1000 breast tumours with a large immunohistochemical panel of 25 antibodies. These included antibodies for the luminal phenotype based upon the expression of cytokeratins 7/8, 18 & 19 and the basal phenotype with cytokeratins 5/6 and 14. In addition to this hormonal markers ER, PR, androgen receptor (AR), tumour suppressor genes (p53, BRCA1, Anti-FHIT), cell adhesion molecules (E-cadherin, P-

cadherin), mucins (Muc-1, Muc-1 core, Muc-2), apocrine differentiation (Anti-GCDFP-15), and neuroendocrine differentiation (chromogranin A, synaptophysin) were evaluated. Using an agglomerative clustering approach, Abd El-Rehim et al classified tumours into six groups, with nine of the twenty five antibodies being key discriminatory markers (AR, c-erbB-2, CK18, MUC1, CK5/6, p53, nuclear BRCA1, ER and E-cadherin). Interestingly, the results broadly correlate with the genomic arrays of Perou and Sorlie and defined six classes of breast cancer. Two were hormonal positive luminal groups having similar immunohistochemical profiles to luminal A and B sub-groups but having differing overall disease free survival, a third group had c-erbB2 (HER2) positivity with strong Muc1 expression, altered E-cadherin and weak hormonal marker expression. The fifth group showed elevated p53 and basal marker expression and lower hormone receptor expression, group six consisted of HER2 and E-cadherin positive tumours that had weak or no hormone receptor status and MUC1. Group four only consisted of four tumours so it is difficult to draw conclusions for this division.

These authors similarly showed, like Perou et al, that the IHC identification of sub-groups was of prognostic value in invasive breast cancer. The grouping determined in this series was defined by a panel of antibodies reflecting the various genetic pathways that may be responsible for tumour growth. Other workers, such as Nielsen (82) and Rakha (84), have also demonstrated that the basal subgroup can be identified using IHC markers.

In 2011, the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer agreed to define subgroups for invasive breast cancer based on clinicopathological criteria, including immunohistochemistry for ER, PR, HER2 and Ki67, as opposed to a gene expression array criterion. Their reasoning was that systemic therapy approximately matches the subtype classification (97). They proposed the following subgroups: luminal A (ER and/or PR positive, HER2 negative and Ki67 <14%), luminal B/HER2- (ER and/or PR positive, HER2 negative and Ki67

≥14%), luminal B/HER2+ (ER and/or PR positive, HER2 positive), HER2+/ER- (non luminal) (ER and PR negative and HER2 positive) and triple negative (ductal), (ER, PR and HER2 negative). To date only one study (98) has used this specific criteria to assess DCIS, finding no link between the subgrouping and risk of recurrence, as either invasive carcinoma or DCIS, over a ten year period.

1.8 Conclusion

DCIS still represents a clinical dilemma for patient management. The inability to reliably predict possible progression to invasive breast cancer results in the possible overtreatment of what can be regarded as a non-life threatening condition. Current treatment options remain controversial with no consensus upon the choice of mastectomy, or breast conserving surgery with or without radiotherapy, or indeed adjuvant hormone treatments.

The natural progression and molecular pathology of DCIS also still remains poorly understood. This thesis aims to address some of these problems by analysing a series of DCIS cases with known clinical outcome. This includes DCIS lesions that have not progressed (pure DCIS), DCIS lesion that have had a DCIS recurrence and DCIS lesions that have an invasive recurrence with or without associated DCIS. A further series of cases of pure DCIS and DCIS admixed with invasive disease contemporaneously allows comparison of biomarker and genomic changes in the different disease states. Analysis of genomic changes using Molecular Inversion Probe arrays (MIPs) and immunohistochemical (IHC) profiles in these cases will hopefully identify possible biomarkers of interest and genetic abnormalities that confer greater or lesser risk of progression from DCIS to invasive disease and allow comparison of groups of lesions with regard to genetic changes and biology.

1.9 Hypothesis

An immunohistochemical and genomic assessment of “pure” DCIS and of DCIS associated with invasive breast cancer will identify differences in biomarkers and gene expression between individual cases. Within immunophenotypic sub-groups variations in biomarkers and gene patterns may identify cases with increased risk of recurrence and/or progression.

Summary of Aims:

1. To determine the expression of a range of immunohistochemical markers in DCIS lesions and determine if these markers can identify molecular subtypes similar to those found in invasive breast cancer.
2. To create a database of significant size providing information on DCIS samples with genomic, immunohistochemical and demographic data. This database will provide a resource for further research beyond the scope of this thesis.
3. To undertake an analysis of archival DCIS samples to identify genomic variation, abnormalities and changes between different pure DCIS lesions and DCIS lesions associated with invasion
4. To identify any protein and/or genomic changes in all DCIS (pure and that associated with invasion) compared to invasive breast disease to determine potential biomarkers responsible for progression to an invasive state.

Chapter 2. Construction of Tissue Microarrays (TMAs) for Immunohistochemical Evaluation of Ductal Carcinoma in Situ

2.1 Introduction

In this study, the evaluation of multiple immunohistochemical markers is used to determine if DCIS can be classified and sub-typed in a similar fashion to invasive breast carcinoma on an immunohistochemical level. In order to assess multiple markers on significant numbers of DCIS cases, TMAs have been used. The nature of DCIS, including size and distribution of lesions and intralesional heterogeneity, introduces some challenges in TMA construction. The design and manufacture of TMAs had to address several criteria specific to DCIS, as well as some general principles not often cited in the literature. These include core size, number of cores, core placement, tissue selection criteria, heterogeneity of tissue and subsequent TMA analysis. These issues are discussed below.

2.2 Materials and Methods

A search was carried out for DCIS specimens within the King's Health Partners Tissue Bank (KHPB), formerly the Guy's and St Thomas's Breast Tissue and Data bank (BTBD), yielding a list of 579 cases for evaluation. The specimens were defined as containing a "pure" DCIS component not associated with invasive carcinoma, or DCIS as a recurrence. All tissue had been uniformly fixed in 10% neutral buffered formalin within 60 minutes of surgery. The haematoxylin and eosin (H&E) stained slides for cases were retrieved from the archive and screened by a consultant breast histopathologist (SEP) to reclassify each lesion using the current classification and

terminology. Each case was then assessed for suitability for coring for tissue microarray. Cases where core biopsy material alone was available were discounted due to lack of tissue.

Once the slides had been reviewed, the matching formalin fixed paraffin wax embedded (FFPE) blocks corresponding to the marked slides were retrieved from the archive. The blocks were assessed for amount of remaining tissue and any deemed insufficient were excluded from the study. It was noted that for some of the cases the tissue in the FFPE blocks did not match the shape of the tissue sections due to taking of multiple sections for other, previous, studies. All the blocks were subsequently re-embedded and a single section cut and H&E stained for marking using an optical marker (Nikon-Melville, New York U.S.A.) attached to a microscope (Figure 14).

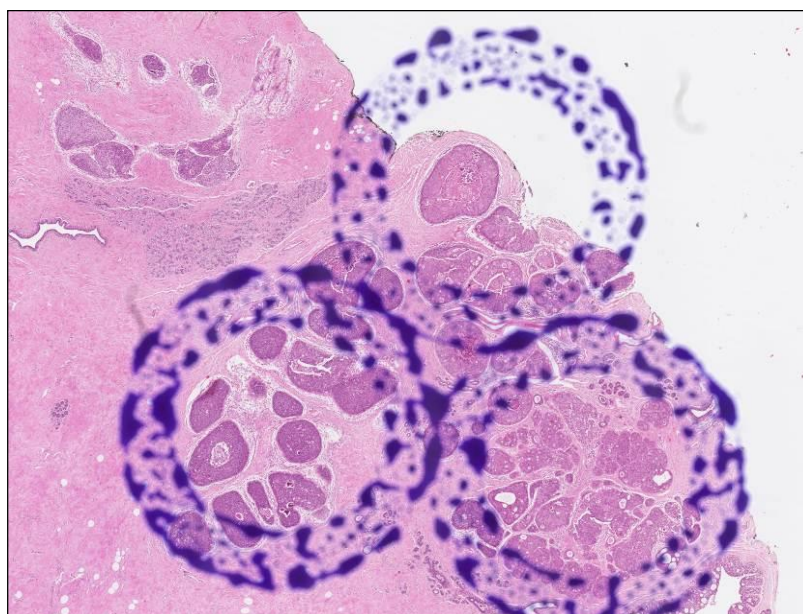


Figure 14. Optical marking of disease on an H&E section.

Once areas of interest had been marked, TMAs were constructed. As described in Chapter 1, the basic principle is to take an empty paraffin wax block (recipient block) with which a small hole is drilled using a first stylet. Then a second, slightly larger, stylet is used to sample the paraffin wax block containing the tissue of interest (donor block). This core of tissue is then inserted into the small hole in the recipient block.

Once all the desired cores have been transferred from the donor to the recipient block then the newly constructed TMA undergoes an annealing process. This process involves heating and cooling the newly constructed TMA block to just below the melting temperature of the wax to enable the cores to meld to the wax of the recipient block. Various temperatures and times are used dependent upon core size and tissue type. For this study, trial and error demonstrated that placing the new TMA blocks in a pre-warmed oven at 56°C for ten minutes followed by 30 minutes cooling repeated thrice produced TMAs with suitable annealing to prevent core loss.

Each case of DCIS had three 2mm cores taken, with each of these cores placed into a separate block labelled A, B and C respectively. The standard size for TMA cores in studies of invasive carcinoma is 0.6-1.0 mm (99). A larger core size (2mm) has been used in this study to be certain of obtaining representative morphology from ducts bearing in-situ carcinoma. This results in a larger tissue surface area to demonstrate IHC expression and to assess the presence of staining heterogeneity. Each TMA block consisted of twenty DCIS test cases, and five control tissue cores of either kidney or liver tissue, to enable easy orientation of cores when examining the sections under light microscopy (Figure 15).

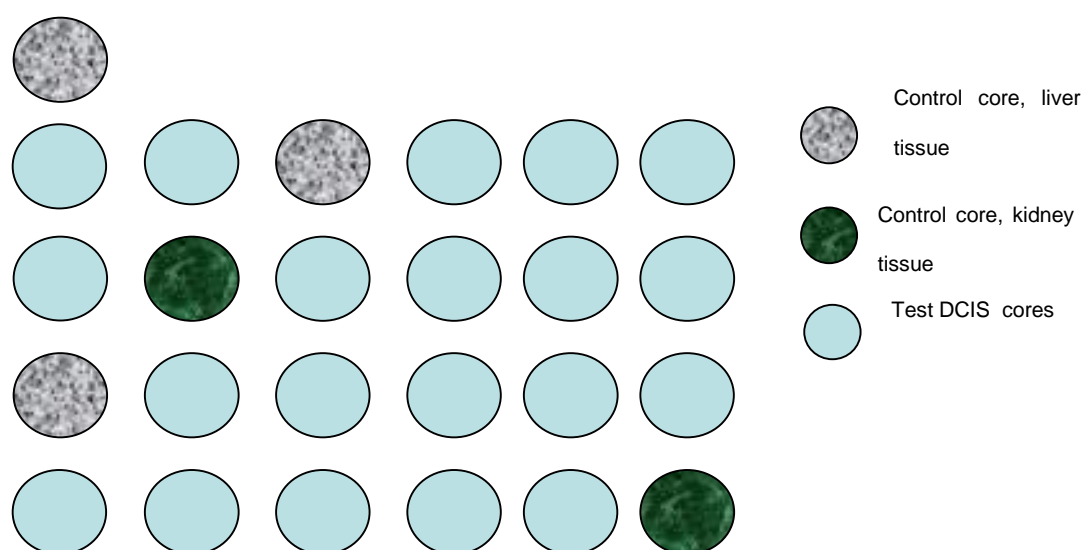


Figure 15. Template for TMA composed of 2.0mm cores

Once constructed, a peripheral surface section was cut from the TMAs and an H&E stained slide produced in order to assess the accuracy of the coring and confirm the presence of DCIS. Examples of sample 2.0mm cores with DCIS present are shown in Figure 16 with an example of a whole TMA section in Figure 17.

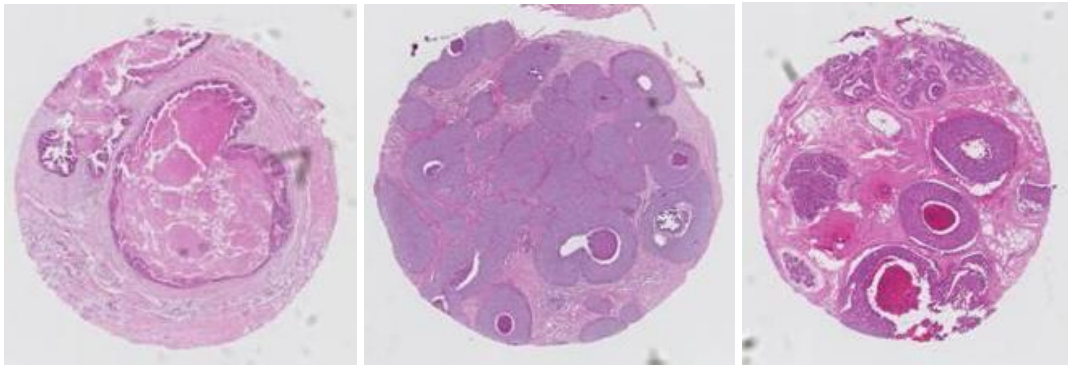


Figure 16: DCIS in 2.0mm cores

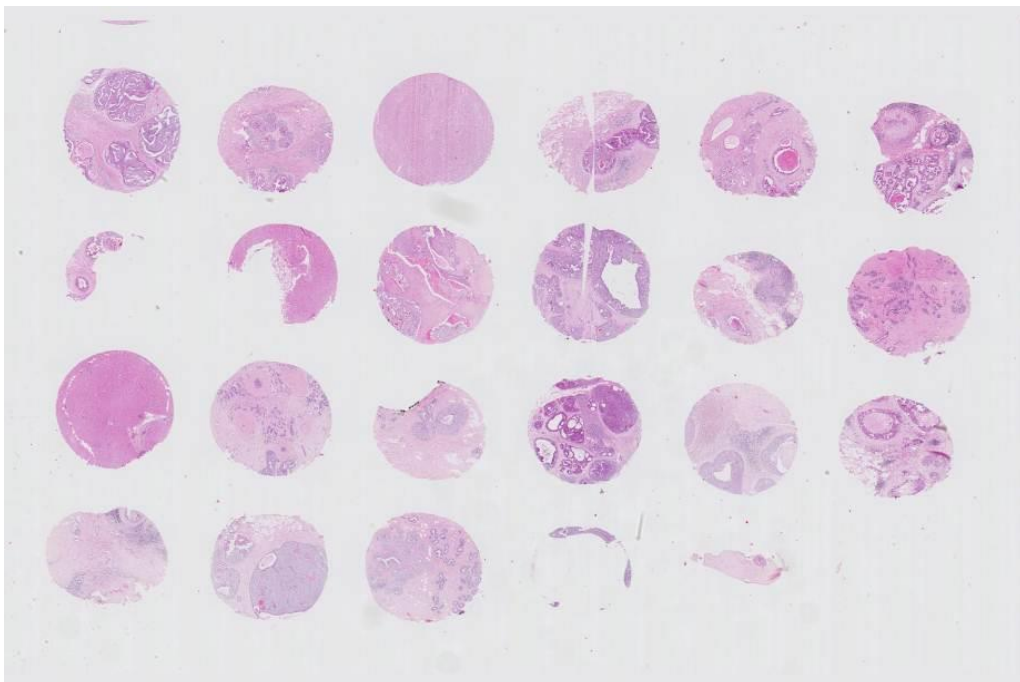


Figure 17: TMA construct

Use of tissues and data (see 3.2.1) was approved under NHS REC agreement (07/H0874/131) with Biobank Access Committee approval.

2.3 Results

2.3.1 TMA Constructs

The 579 cases of DCIS in the initial database yielded a total of 280 cases of pure DCIS suitable for TMA construction. These were made into 14 tissue microarray blocks in triplicate giving a total of 740 cores for IHC staining.

2.3.2 Scanning

All TMAs had a 4µm section cut and stained with H&E to verify the presence of DCIS within cores. A total of 42 H&E sections were taken and scanned using a “Nanzoomer” digital scanner (Hamamatsu, Welwyn Garden City, UK) to provide digital images for analysis. Loss of DCIS due to either missing cores, incorrect coring from donor block or lack of viable cells was evident in 15-20% of cases.

2.4 Discussion

2.4.1 Tissue Selection Criteria

The selection of good quality tissue for inclusion in any TMA is paramount. However, the vast majority of TMAs are drawn from archival diagnostic stores with very few having tissue prospectively acquired. A major problem with the use of archival tissue is that fixation and processing is typically poorly documented, and often highly variable even within a single institution. The archive based at Guy's and Thomas' Tissue and Data bank has specimens dating back to the early 1970's available for research. The archive also includes accurate records of time to fixation and duration in fixative; the breast tissue used in this study were expertly sliced and placed in

formalin within 30 minutes of surgery. The maximum duration of fixation was 48 hours. The quality of the archive owes much to the foresight of the individuals who established the tissue bank, and who recognised the importance of adequate fixation and detailed record keeping before the advent of modern immunohistochemistry and molecular techniques.

2.4.2 TMA cores size, number and distribution within a TMA block

There is variation in the literature as to what is considered “representative” when selecting the number, size and distribution of cores within a TMA. Core diameter and number ranging from a single 1.0 mm core to six 0.6 mm cores have been suggested as suitable. Some authors have reported that a single core, regardless of size (1.0mm to 0.6mm), correlated with whole section mounts when assessed for ER, PR, HER2 and Ki67 (100-102). However, others have advocated the use of multiple cores to provide more accurate concordance; Graham et al suggested that a minimum of four assessable cores from six sampled areas were required for accurate evaluation of HER2 status of invasive breast cancer (103). Bhargava showed that between two to four cores had a high (99%) concordance with whole sections for HER2 status by fluorescent in-situ hybridisation (FISH), again in invasive disease (104). Sapino et al recommended that at least 4 cores are required and that where a negative hormonal marker such as ER or PR is found then the IHC should be repeated on a whole section (105). Thus there is little agreement, for even standard biomarkers, on the number of TMA cores that should be assessed, let alone the number of cells required.

Table 1 compares the surface areas of the most common cores sizes: 0.6mm, 1.0mm, 1.5mm and 2.0mm and the total surface area obtained when multiple cores are examined. Taking one 1.0mm core is roughly equivalent to three 0.6mm cores. This excludes demonstration of regional heterogeneity within a cancer, but for markers with uniform expression significant time and cost savings could be made by use of a larger core size. Where multiple cores are taken from lesions of known heterogeneity,

increasing the core size from 0.6mm to 1.0 mm will give a threefold increase in tissue area for analysis.

Core Diameter (mm)	Core Radius (mm)	Core Circumference (mm)	Core Area (mm ²)
0.60	0.30	1.885	0.283
1.00	0.50	3.142	0.785
1.50	0.75	4.712	1.767
2.00	1.00	6.283	3.142

Table 1: Surface areas of different TMA core sizes.

There is an inevitable loss of cores when cutting TMA sections and performing subsequent tests such as IHC. Such loss can be due to tissue type, fixation, poor region selection, poor annealing of cores in the donor block, core size, section thickness, section drying and overuse of organic solvents. Technical expertise can minimise some of these, although the use of archival material that is poorly processed will have an influence on core preservation. Heating of sections for processes such as antigen retrieval in IHC and denaturation of DNA in ISH are a particular cause of core loss and a percentage will be lost in such processes, even when performed by expert technicians.

Histopathological sections of diseased tissue are, by their very nature, a small representation of an entire tumour or disease state. Rarely is a tumour sampled in its entirety with every section of tissue cut and examined for tumour heterogeneity with repeated antibody testing. Large tumours are excised and the pathologist selects representative pieces which, in general, have only one H&E section taken and examined from the block. TMAs have, since their inception, have been dogged with the criticism that they are not representative of the entire tumour and cannot capture

the heterogeneity seen in whole sections. Camp et al (99) point out that the general consensus is that in most cases two 0.6mm cores are representative of a whole section but that some antibodies, such as those involved in hypoxia, that show heterogeneous expression may require more cores. In this study, selection was based on the diagnosis of DCIS. Larger cores were chosen as they allowed entire duct spaces, with their surrounding myoepithelial layer and basement membrane to be visualised within the TMA core. This would also address issues of potential staining heterogeneity (by manufacturing triplicate blocks) and included a sufficient quantity of tissue to be considered representative of a lesion when compared to the guidelines quoted in the literature. Indeed the actual sampling in this thesis exceeds the vast majority of current studies. It is anticipated that this would give a true reflection of IHC staining patterns in the tissues examined.

Chapter 3. Immunohistochemical Studies on Ductal Carcinoma in Situ

3.1 Introduction

The expression of key proteins, or biomarkers, is of prognostic and/or predictive importance in invasive breast cancer. Panels of proteins have been proposed that can divide breast cancer into clinically important subtypes with differing survival outcomes. However, there is currently a lack of information as to whether these tumour subtypes are represented at the in situ stage, and their prognostic significance in DCIS. Immunohistochemical experiments looking at a number of proteins of known importance in invasive breast cancer were carried out for this thesis to investigate their expression within DCIS. Sections 3.3 to 3.9 refer to specific types of antibody with regard to their target antigen and the rationale for inclusion into a prospective panel. The procedure for staining for all antibodies was similar and is summarised below with details of any variations given.

3.2 Material and Methods

3.2.1 Correlation of Immunohistochemical Results to Patient Data and Outcome

See Appendices 2, 3 & 4

In order to evaluate the effect of any IHC biomarker upon the likelihood of recurrence, whether invasive, in situ or not present, it was necessary to determine the patient outcome in our cohort. Patient demographics including age, treatment, details of local, ipsilateral and contralateral recurrence, if any, were retrieved from the King's Health Partner's Breast Cancer Biobank records.

Samples of pure DCIS tissues from this cohort of patients were taken between the years 1974 and 2005. Patient details in some instances were incomplete and not

suitable for further analysis. Reasons for these losses included patient relocation, loss of records or conflicting data. This reduced the total number of TMA cases available for correlation with patient demographics. For these cases TMA core loss also occurred in a small percentage of cases giving a slight variance in numbers available for each antibody assessed.

The data in the bank is prospectively populated with follow up data from clinical visits by patients and from cause of death registration. Patients were diagnosed with DCIS between 1974 and 2005.

Table 2 summarises patient demographics. Records for 180 patients were available. Mean patient age at diagnosis was 72 (range 22-90).

Records list 112 patients with an excision biopsy as first form of surgery, 11 had microductectomy, 2 had a needle localisation excision biopsy, 5 had a wide local excision and 12 had a simple mastectomy. Data on type of first surgery was unavailable for 20 patients. The majority of patients underwent further treatment (n=131), of these patients 41 had a simple mastectomy, 50 had a wide local excision, 7 were listed as both wide local excision and simple mastectomy, 14 had a modified radical mastectomy, 2 had an excision biopsy, one biopsy and one microdochectomy. There was no data on the type of subsequent surgery for two patients.

Twenty three patients had a known recurrence more than 3 months (range 108-4363 days) after original diagnosis and surgery.

3.2.2 Histological Assessment of Necrosis and Inflammatory Response.

Histological re-assessment of samples (Chapter 2) provided details of the degree of comedo-type necrosis and of inflammatory response as well as cytonuclear grade and architecture of the DCIS. The histological data frequencies are shown in Table 2. The presence of necrosis was scored as “mild”, “moderate” or “marked” where 10% or less of the ducts bearing comedo-type necrosis were defined as a mild degree,

moderate corresponded to approximately 10-50% and marked to 50% or more of ducts exhibiting necrosis (46). Assessment of the presence of inflammation was based on 'routine' categorisation of the degree of chronic inflammatory cell infiltrate adjacent to and/or surrounding the duct spaces bearing DCIS. Previous work in classifying inflammation in DCIS has shown good reproducibility between consultant pathologists in (RM and SEP unpublished data).

Table 2: Patient Demographics and Histological Details.

		Number of patients	% (of 180 in TMA series)
Mean patient age at diagnosis (range)	72 (range 22-90)	166	92
Size of DCIS (mm)	Average 14.9 Median 14.9 Range 1.0 - 31.0	168	93
Dominant architecture n (%)	Solid 74 (43) Cribriform 69 (40) Micropapillary 18 (10) Papillary 13 (7)	174	97
Side n (%)	Left 92 (58) Right 68 (42)	160	89
Cytonuclear grade n (%)	High 103 (60) Intermediate 57 (33) Low 13 (7)	173	96
Invasive Recurrence n (%)	Yes 18 (13) No 140 (87)	161	91
Invasive Recurrence Site n (%)	Ipsilateral 9 (50) Bilateral 1 (6) Unknown side 8 (44)	18	10
Invasive Recurrence Grade n (%)	High 8 (77) Intermediate 3 (23) Low 0 (0)	11	7
DCIS Recurrence	2 (100)	2	1
DCIS Recurrence Site	Ipsilateral 1 (50) Unknown side 1 (50)	2	0.5 0.5

DCIS Recurrence Grade	High	2 (100)	2	1
New Lesion (contralateral)		3	3	2

3.2.3 Immunohistochemical Experiments

Tissue sections from the manufactured in-house TMAs (Chapter 2) were cut at 4µm and dried overnight at 37 °C before being baked at 57 °C for two hours prior to IHC staining. Immunohistochemical staining, using antibodies raised against oestrogen receptor (ER) clone 6F11 (Leica, Newcastle, UK), progesterone receptor (PR) clone 636 (Dako, Ely, UK), HER2 (Oracle kit, Leica, Newcastle, UK) cytokeratin 5/6 (CK5/6) clone D5/16 B4 (Dako, Ely, UK), cytokeratin 5 (CK5) clone XM26 (Leica, Newcastle, UK) , cytokeratin 14 (CK14) clone NCL-CK14 (Leica, Newcastle, UK), Ki67 (clone MIB1) (Dako, Ely, UK), MCM2 –clone NCL-MCM2 (Leica, Newcastle ,UK) and epidermal growth factor receptor (EGFR) clone 384 (Leica, Newcastle, UK), was performed. All TMAs were stained in triplicate to reduce the loss of data due to core loss and ensure sufficient cells were available for assessment. Sections were stained using a two-step compact polymer chain biotin free IHC protocol using a 3-3'diaminobenzindine chromogen on the BondMax™ Leica automated staining system. HER2 tests were carried out on the same system using the HER2 Oracle kit (Leica, Newcastle, UK). All slides underwent blocking for endogenous peroxidase and a haematoxylin counterstain prior to coverslipping in an organic medium. Antigen retrieval was undertaken for each antibody according to manufacturer's recommendations. All slides were stained alongside appropriate known positive control sections for quality control purposes. ER and PR status was previously recorded for some of the cases in this study, however due to changes in IHC methodology and improvements in antibody clones it was determined that repeat would be undertaken for all. Appendix 4 lists reagents and protocols for immunohistochemical staining.

3.3 Hormone Receptors

3.3.1 Oestrogen Receptor Structure and Function

Oestrogens are steroidal sex hormones involved in a large number of physiological processes (106). They have a major effect on the female reproductive system, including ovulation, pregnancy, childbirth and lactation (106). They are also involved in the cardiovascular system, immune response, central nervous system and bone growth (106). Many of the physiological actions of oestrogen are mediated via binding to the oestrogen receptor (ER) (106-108).

ER is a ligand-activated intracellular transcription factor responsible for the biological effects of oestrogen via gene regulatory pathways (109-111). ER belongs to a groups of proteins known as the nuclear hormone receptor (NHR) superfamily which includes progesterone (PR), testosterone and vitamins A and D (111). There are two identified isoforms, ER α and ER β , coded for by the genes ESR1 (located 6q24-q27) and ESR2 (located 14q21-q22). These isoforms have 95% amino acid homology (110). ER protein structure (Figure 18) consists of a variable N terminal domain containing an activation domain, AF1, followed by a DNA binding domain (DBD). The DBD is linked to a variable hinged region connected to a ligand binding domain (LBD). The LBD contains a second activation domain AF2. The activation domains regulate the transcriptional activity of ER (110, 112). The DBD allows the ER to bind to specific DNA loci within the nucleus, known as oestrogen response elements (EREs), due to the presence of two highly conserved zinc fingers distinct to the NHR superfamily. The variable hinged region of ER is a flexible region of the protein and acts as a

nuclear localisation signal connecting the DBD and LBD (110). The LBD acts as a “molecular switch” which, upon binding the hormone oestrogen, becomes transcriptionally active (112, 113). ER α and ER β binding of 17 β estradiol (E2) via the EREs and the transcriptional activation has been designated the classic pathway of oestrogen action (114). The binding of E2 to the receptor results in conformational changes leading to dimerisation with coactivator or corepressor molecules which regulate downstream physiological responses (115).



Figure 18: Structure and Functional Domains of ER

3.3.2 Oestrogen Receptor and Breast Cancer

Unlike most of the NHR superfamily, oestrogen receptor (ER) (and androgen receptor (AR)) have the ability to stimulate cellular proliferation and are essential for the normal development of the female breast during puberty and pregnancy (116). However, it is this very ability that, when aberrations occur, can lead to tumourigenesis and uncontrolled proliferation. The exact mechanism of how the ER/E2 complex initiates cellular proliferation remains elusive.

IHC expression of ER status is a weak prognostic marker of independent clinical outcome but a strong predictive marker of response to therapy in invasive breast cancer (117). Approximately 60-70% of breast cancers in women under 50 years and 80% of breast cancers in women over 50 are ER positive (118). Unlike many IHC markers, ER is used clinically to help determine the likelihood of response to endocrine therapy for patients with invasive breast cancer, which can be in the form of antagonists that directly block the receptor pathway (Tamoxifen) or indirect blocking such as the aromatase inhibitors which prevent conversion of androgens to oestrogen (119).

3.3.3 *The Oestrogen Receptor and Ductal Carcinoma in Situ*

The expression of oestrogen receptor in DCIS may have important implications for the treatment of precursor lesions, with some studies showing ER, PR and HER2 predict an increased risk of local recurrence (120, 121). Up to 80% of DCIS lesions are ER positive. (122). However, treatment of DCIS with anti-oestrogen therapy such as Tamoxifen has not shown an overall decrease in mortality (54). Randomised trials by the National Surgical Adjuvant Breast and Bowel Project (NSABP) have demonstrated a reduction of incidence of both invasive breast cancer and DCIS with Tamoxifen treatment for women aged over 35 recognised to be at increased risk of developing breast cancer (123).

3.3.4 *Progesterone Receptor*

Progesterone receptor (PR) is another member of the steroid receptor superfamily important for growth and development in the breast and female genital tract. PR expression is induced by oestrogen, and is thought to be an indicator of an intact ER pathway. Thus, PR positivity may be an adjunct to prediction of response to oestrogen related therapies (118). PR is also often advocated as an IHC marker where a possible false negative ER result is suspected (118). It has been shown that some ER+ PR- invasive breast tumours have a reduced response to tamoxifen compared to ER+ PR+ tumours, however the exact mechanism as to how both oestrogen and progesterone interact is not known (118, 119, 124). Tamoxifen therapy induces a greater drop in PR than ER levels, with some lesions losing PR expression completely; this may be one mechanism in the development of acquired tamoxifen resistance and a poorer overall survival (OS) in patients with invasive breast cancer (124). The ATAC trial (125) studied the response of ER positive breast disease treated with tamoxifen or anastrozole (arimidex) alone or in combination. Results from this study showed that patients treated with anastrozole had prolonged disease free survival, time to recurrence, distant metastases and contralateral breast cancers after

five years compared to tamoxifen (125) . These findings were also reported in a long term follow up of the ATC trial after 10 years (126).

3.4 Epidermal Growth Factor Receptors

3.4.1 *Epidermal Growth Factor Receptor Structure and Function.*

The epidermal growth factor receptor (EGFR) or ErbB family of tyrosine kinases (phosphorylation enzymes) comprises four related transmembrane proteins (HER1 to HER4) that can form oligomers, usually in the form of a homo-dimers or hetero-dimers (127). The HER family of proteins acts as conduits for the activation of intracellular signalling pathways from extracellular signals (128). All four of the HER proteins are characterised by an extracellular ligand binding domain, a transmembrane domain and an intracellular tyrosine kinase domain (128). Crystallographic studies by Burgess et al (129) defined the structure and ligand induced dimerization, however the exact kinetic mechanisms remained elusive. Zhang et al (130) elucidated, through further crystallographic & mutational analysis, that the allosteric mechanism was similar to CDK2/cyclin A activation, where the C-terminal lobe of one kinase interacts with the N terminal of the adjacent kinase. The Erb family of proteins plays key roles in mediating the cellular proliferation of cells, differentiation, metabolism and motility (131).

The *HER2* gene was first identified in human breast cancer by King et al in 1985. Also known as ErbB2 (referring to both human and rodent gene), "*HER2*" refers to the human gene and protein and not the rodent (128). Unlike HER1, HER3 and HER4, HER2 does not bind specific ligands but relies on hetero-dimerisation, with one of the other ligand-bound HER proteins (132) (Figure 19).

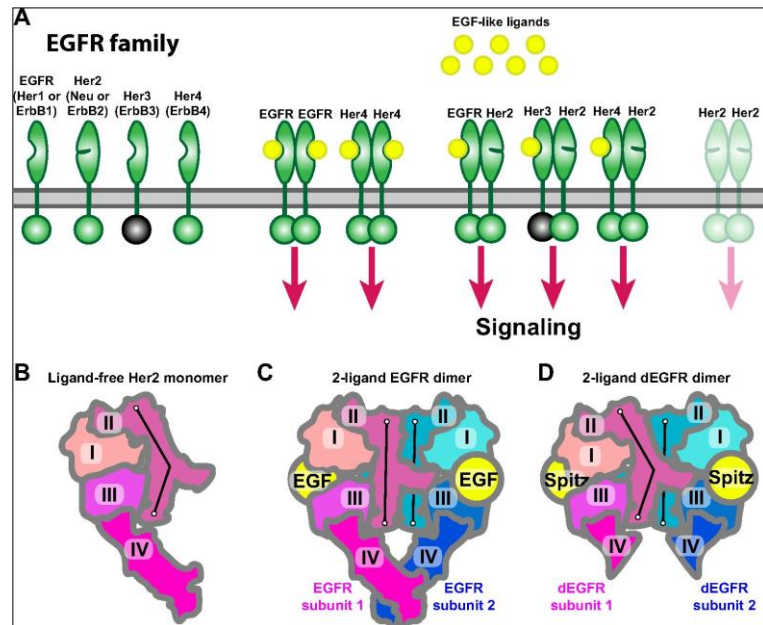


Figure 19: EGFR Structure. DOI: <http://dx.doi.org/10.7554/eLife.00708.003>
reproduced through Creative Commons Attribution license: original image from
reference (132)

The overexpression of both HER1 (EGFR) and HER2 (C-erbB2) has been implicated in many solid tumours, including bladder, breast, colon and lung cancers (127, 133, 134).

EGFR (HER1) has been shown to be expressed in the basal-like invasive breast tumours and in some DCIS (127, 134) and is one of the targets currently under investigation for monoclonal antibody therapy with drugs such as erlotinib, cetuximab, and lapatanib in invasive breast cancer (135, 136). There are several antibodies commercially available for the immunohistochemical assessment of EGFR. Although membrane reactivity is regarded as 'positive' for EGFR, significant staining variation between antibodies exists (unpublished data) and there is yet to be a defined pathological scoring system such as the HER2 scoring system for this antibody.

In breast cancer HER2 overexpression is present in between 15-25% of invasive lesions. Although associated with a poorer prognosis overall, targeted therapy, for

example with the drug trastuzumab (Herceptin), is available (137) and widely recommended and effective for patients with HER2 positive invasive breast cancer.

3.5 Proliferation Markers

3.5.1 *Introduction*

Deregulated proliferation is one of the hallmarks of the cancerous state (138). In order for a tumour to grow and infiltrate surrounding tissues, the ability for a neoplastic cell to replicate is essential for the progression of a lesion. Several IHC markers have been established as indicators of proliferation, and have found use in both the research and diagnostic laboratory setting. In this study, Ki67 and the minichromosome maintenance protein 2 (MCM2) have been selected as suitable indicators of proliferation.

3.5.2 *Ki-67*

Ki-67 is an established marker for the assessment of cellular proliferation (139). First used to show proliferative activity in frozen or freeze-dried sections of human lung tissue (140) it is one of the most utilised biomarkers in cancer research. A Pubmed enquiry in 2015 using the search criteria “Ki-67 & cancer” yields over 15,000 articles, a search for Ki-67 alone yields over 19,000. The large number of articles belies the fact that the exact function of this protein is still not fully understood and there is a relative paucity of papers looking at Ki-67 structure and function (139). However, it is almost universally accepted that Ki-67 protein expression is essential for the cell cycle (139). Removal of phosphorylated Ki67 from cells using antisense nucleotides prevents cell cycling (141). Loss or inhibition of Ki67 results in ribosomal RNA synthesis inhibition (142). Ki67 is expressed during all stages of the cell cycle, G1, G2, S and mitosis, however, it is not expressed in G0 (143).

3.5.2.1 *Ki-67 nomenclature*

The Ki-67 antigen is known to be a large protein and is referred to as Ki67 protein, as opposed to the original antibody, raised against the protein which is referred to as the Ki67 antibody. This antibody was the prototype for other Ki67 antibodies raised to different epitopes on this large molecule and was produced by immunizing mice with nuclei from a Hodgkin lymphoma cell line (L428) (139, 141). These include the clone MIB1 (Molecular Immunology Borstel), MIB5 and TEC-3. The name Ki-67 originates from the city of discovery (Kiel) and the position in a 96 well plate of the original clone (139).

3.5.2.2 *The Structure and Function of Ki-67*

Ki-67 has 2 prominent forms detectable at the protein level (Figure 20), a 395kDa and a 320kDa protein encoded by nearly 30,00 base pairs (141). Schultze et al (144) published the cDNA sequence encoding for the *Ki-67* protein in 1993 identifying 2 spliced isoforms of mRNA (modifications of the RNA transcript by removing introns and introducing exons) which differed by the absence in one region encoded by exon 7. They also describe a series of 366 base pair “Ki67 repeat” elements in a 6,845 base pair exon located centrally within the antigen. This exon (138) is one of the largest mammalian exons known. They concluded that Ki-67 has no known homology with other known amino acid sequences and it is a short-lived category of cell cycle associated nuclear non-histone protein. Further studies by Hofmann and Butcher identified a forkhead associated domain (FHA) in Ki67 as well as other unrelated proteins found in yeast, bacteria and mammals but all believed to be involved in DNA repair, synthesis or cell cycle regulation (145). The Ki-67 protein contains 143 protein kinase C, 89 casein kinase II and 2 tyrosine kinases (protein kinases phosphorylate serine, threonine, tyrosine, or histidine residues to modify a protein function) (139). Analysis of the amino acid sequence has led to the prediction of ten nuclear targeting

signals, which may explain why Ki-67 is only found in the nuclei during interphase (139, 146-148).

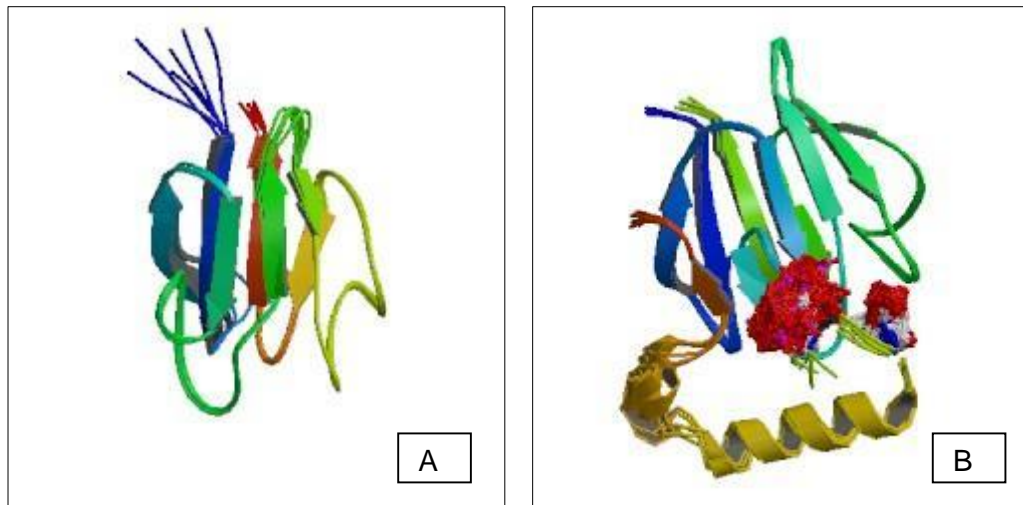


Figure 20: Ki-67: Protein structure 1R-21-1-120 (human) A & 2AFF-1-120 (human) B. Images courtesy of RCSB PDB rendering based on 1r21.en.wikipedia.org

Despite the fact that Ki-67 has been described, and used in immunohistochemical studies, for over thirty years the exact function remains elusive as do its possible interactions with other cell cycle markers. There are several possible explanations why Ki67 function is difficult to elucidate. Scholzen et al suggest that its large size and susceptibility to proteolytic cleavage make biochemical assays difficult. Secondly, they cite the lack of similarity to other proteins makes study of this molecule with standard methods difficult. To date there is no definitive explanation of Ki-67 function. Although believed to be expressed solely in the active proliferative phases of cell cycling (G_1 , G_2 , S and mitoses) there is some evidence that Ki67 can be detected in quiescent cells (G_0) and has a possible role in RNA transcription (149) and RNA synthesis (142).

3.5.2.3 Ki-67 as a Biological Marker in Cancer

There have been many studies showing that Ki67 assessment has prognostic value in invasive cancer. Brown and Gatter's review in 2002 showed a total of 23 papers

examining Ki67 in breast cancer alone as having an independent prognostic value (141). A Pubmed search on the 26th April 2015 for “Ki-67, breast cancer and independent prognostic factor” yielded over 160 journal articles citing Ki67 alone, or in combination with other markers, having potential value as a prognostic marker. More recently, the focus has been on the use of Ki67 as a predictive factor (i.e. in selection of response to therapy) with Darb-Esfahani al reporting that Ki67 can be utilised as an independent predictor of response to anthracycline/taxane-based neoadjuvant chemotherapy regardless of invasive tumour type (150) although Jones et al found no association with Ki67 and clinical response (151). Other studies have reported that changes in Ki67 level are a valuable predictor of response to hormone therapy. Finally, groups, such as Cheang et al, have used a Ki67 Index in association HER2 and ER to define/classify the luminal B subgroup of invasive breast cancer using immunohistochemistry (152).

3.5.2.4 Ki67 and Ductal Carcinoma in Situ

In DCIS, Ki67 expression in conjunction with COX-2 and p16 expression, has been suggested to be an important factor and associated with an increased risk of developing subsequent invasive disease (120). Altinas et al examined a range of biological markers, including Ki67, as possible predictors of local recurrence and no found evidence that Ki67 alone or in conjunction with other markers was significant (153).

However, much of the data on Ki67 and definition of sub-groups for comparison with outcome has been undertaken on invasive disease. Zhou et al used Ki67 with a cut-off of >14% to distinguish positive cases to assign subgroups in invasive disease. Those with >14% were defined luminal A and those with <14% were defined as luminal B/HER2 negative (as defined by the St Gallen International Expert Consensus from Cheang's publication (145), although they found no significant prognostic value in their cohort (97). Ruiz et al (154) compared three molecular classification

subgroups, finding that only the classification including Ki67 as a marker improved prediction of overall survival in invasive disease.

3.5.3 Mini Chromosome Maintenance Proteins

3.5.3.1 The Structure and Function of Mini Chromosome Maintenance Proteins

Unlike Ki67, the structure and function of the Mini Chromosome Maintenance (MCM) proteins are well understood. A review by Bochman & Schwacha in 2009 (155) references 280 articles on MCM structure and function in both archaeal and eukaryotic forms. For the purpose of this thesis, information on eukaryotic MCM function is given.

MCM proteins are essential components of DNA replication in the cell cycle (156). Several gene variants exist encoding for the proteins MCM2 to MCM7, which are required to establish a pre-replication complex prior to cellular division. MCM2, MCM3 and MCM5 were first discovered in the yeast *Schizosaccharomyces cerevisiae* (157), followed by the discovery of MCM4, MCM7 and cell division mutations (155). MCM6 was discovered in *Schizosaccharomyces pombe* in 1994 (158). Further studies on *Xenopus Laevis* (African clawed frog) identified that all six proteins (MCM2 to MCM7) were required to couple to DNA for replication to occur (155, 159). In eukaryotic cells, all six cell cycle related MCMs have a similar 250 amino acid region encoding for ATPases. ATPases, also known as AAA proteins, are a diverse group of proteins with a wide range of functions acting as DNA helicases and proteases (155).

MCM proteins form ring-like structures to incorporate ATP binding and enable hydrolysis to alter DNA within the central circle or channel of the ring-like structure and perform helicase activity, unwinding the DNA and forming replication forks (160). The MCM2 to MCM7 helicase structure is essential for initiation and elongation of double stranded DNA (155).

The activation of MCMs from an inactive form to an enzymatic helicase requires the recruitment of several protein cofactors. This includes, but may not be exclusive of, proteins Cdt1, Cdc6, MCM10, geminin and the kinase Cdc7-Dbf4 (161). During G₁ inactive MCM proteins are transported by Cdt1 protein to the origin of replication site where cells undergo "licensing" for replication by a group of six polypeptides, known as the origin recognition complex1-6 (ORC1-6)(162). ORC1-6 is a heterohexameric protein complex which binds proteins Cdt1 and Cdc6 (162, 163) followed by the binding of MCMs thus forming a heterohexamer pre-replication complex (Figure 22). The binding of MCMs requires the presence of ORC1-6 and Cdc6, which provide ATP hydrolysis and thus the energy for this active process (155, 164). Once the MCMs are bound to DNA ORC1-6 and Cdc6 are disassociated from the complex. In S phase of the cell cycle, cyclin dependent kinases (CDKS) promote the formation of DNA replication forks and possibly the initiation of MCM function (155). MCMs demonstrate a helicase activity responsible for the unwinding of the DNA helix prior to replication (165). Once the pre-replication complex has loaded, DNA polymerases are recruited and DNA replication ensues. Sanchez-Berrondo et al (166) demonstrated in *Bacillus cereus* MCMs bind DNA and not only act as a helicase but also function as a primase and a polymerase.

Once replication has occurred, MCMs are inactivated by phosphorylation, leading to their dissociation from the pre-replicative complex thus limiting replication to once per cell cycle (Figure 22).

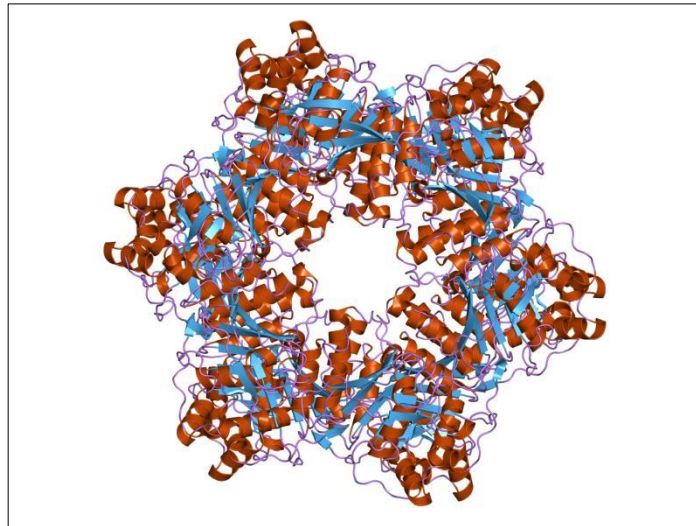


Figure 21: Examples of a Hexameric Helicase Protein Structure for DNA replication.

(Image courtesy of <http://www.ebi.ac.uk/>)

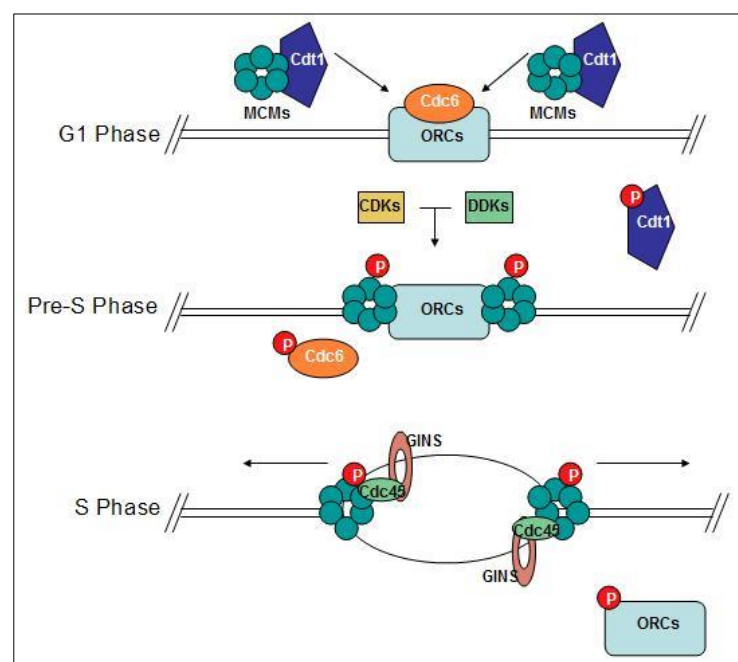


Figure 22: MCM Proteins in the Cell Cycle

(Creative Commons Attribution-Share Alike 3.0 Unported license, author: Azackta1 2013)

3.5.3.2 Minichromosome Maintenance Proteins and Cancer

Previous studies examining MCM2 and MCM5 have shown them as potential alternatives to Ki67 for assessment of proliferation in invasive cancer (167-170). Some have shown MCMs to be independent prognostic markers in invasive breast cancer (171, 172) although MCM expression correlates with Ki67 expression and histological grade (173). The deregulation of MCMs 2-7 has been shown to be an early event in tumourigenesis, both in actively cycling cells and, unlike Ki67, non-proliferating cells that have the potential to enter the cell cycle (167, 170, 174-178). To date, there are only two studies examining the relationship between precursors of invasive breast cancer and MCM2 (179, 180). Reena et al (179) advocate MCM2 as a reliable proliferation marker in DCIS, whilst Nasir et al (180) included MCM2 in a panel alongside TOP2A and BUB1B as “potential molecular biomarkers of malignancy in histologically normal and benign breast tissues”.

3.6 Basal Markers

3.6.1 Cytokeratins 5, 6 and 14

Basal epithelial cells were traditionally described as cells adjacent to the basement membrane (181). However, since Moll's work on cytokeratins (182), the term has become synonymous with those cells expressing the high molecular weight cytokeratins, 5 and 14 and 17 (CK5, CK14, CK17). The term 'basal cell' is often incorrectly applied to myoepithelial cells although the expression of these cytokeratins is not solely confined to myoepithelial cells (183). This has led to a dual meaning for the term basal, which can now refer to either myoepithelial cells or cells expressing “basal” cytokeratins, and is further complicated by the term basal-like breast cancer, which is a molecular phenotype. Basal-like breast tumours are often triple negative (i.e. tumours negative for ER, PR and HER2) and a significant proportion of the triple

negative carcinomas express CK5, 14 or 17. However, the two terms are not synonymous, with some triple negative tumours lacking basal marker expression and occasional basal like tumours being positive for hormone receptors or HER2. Triple negative invasive tumours have a poorer prognosis than many other types of breast carcinoma, with no targeted therapy currently available (86, 184). Thus accurate identification of the premalignant stage could be beneficial in the early identification of these difficult to manage lesions, without the need for adjuvant systemic treatments. In the context of this study, the immunohistochemical expression of one or more of CK5, 5/6 and 14 and / or EGFR is used to define basal breast carcinoma (BBC) as a pathological entity.

3.7 Scoring of Antibody Markers in Immunohistochemical Studies

5.7 TMA slides with immunohistochemical stains were independently scored by a consultant breast pathologist (SEP) and the author (JPB).

Original Excel TMA maps and scores are given in Appendix 2: Excel TMA Maps and IHC Scores

TMA block number	DCIS TMA location	ER Average	PR	HER2	EGFR	CK5/6	CK14	KI67	MCM2
DCIS 15A	A1	0	0	0	0	0	1	1	0

DCIS 15A	A2	0	0	0	0	0	0	x	0
DCIS 15A	A4	2	0	0	0	x	10	1	0
DCIS 15A	A5	0	0	1	3	0	5	40	10
DCIS 15A	B1	x	x	x	0	x	x	x	x
DCIS 15A	B3	0	0	0	3	0	0	10	40
DCIS 15A	B4	0	0	0	0	0	0	30	x
DCIS 15A	B5	2	0	0	1	0	0	15	30
DCIS 15A	C2	5	0	0	0	0	0	5	5
DCIS 15A	C3	1	0	1	2	0	0	5	5
DCIS 15A	C4	5	7	2	0	0	0	1	1
DCIS 15A	C5	8	8	0	0	0	0	0	2
DCIS 15A	D1	x	x	x	x	x	x	x	x
DCIS 15A	D2	x	x	x	x	x	x	x	x
DCIS 15A	D3	0	x	x	x	x	x	x	x
DCIS 15A	D4	x	x	x	x	x	x	x	x
DCIS 15A	E1	2	x	x	x	x	x	x	x
DCIS 15A	E2	0	0	0	0	0	x	0	5
DCIS 15A	E3	5	x	x	x	x	x	x	x
DCIS 15A	E4	8	x	0	0	0	0	1	1
DCIS 16A	A1	0	x	x	x	x	x	x	x
TMA block number	INVASIVE TUMOUR location	ER (Allred)	PR (Allred)	HER2 (HER2)	EGFR (HER2)	CK5/6 (%)	CK14 (%)	KI67 (%)	MCM
DCIS 16A	A2	8	5	x	0	0	0	1	3
DCIS 16A	A4	7	0	3	0	0	0	1	5

DCIS 16A	A5	2	4	2	0	0	0	0	0
DCIS 16A	B1	3	x	x	x	x	x	x	x
DCIS 16A	B3	5	8	2	0	0	0	2	5
DCIS 16A	B4	1	x	x	x	0	0	x	x
DCIS 16A	B5	4	0	0	x	0	0	20	45
DCIS 16A	C2	5	8	2	x	0	0	10	25
DCIS 16A	C3	4	0	1	3	0	0	25	30
DCIS 16A	C4	2	0	0	1	0	0	10	25
DCIS 16A	C5	4	0	3	1	0	0	5	35
DCIS 16A	D1	0	0	0	0	0	0	15	40
DCIS 16A	D2	3	x	x	x	x	x	x	x
DCIS 16A	D3	5	x	x	x	x	x	x	x
DCIS 16A	D4	5	8	0	0	0	0	x	x
DCIS 16A	E1	5	0	x	x	0	0	10	10
DCIS 16A	E2	1	x	3	0	0	0	x	x
DCIS 16A	E3	4	0	3	0	0	0	10	30
DCIS 16A	E4	2	0	2	1	1	0	2	10
DCIS 16A	E5	x	0	x	2	0	0	3	35
INV T 15	A1	0	0	0	3	15	0	8	8
INV T 15	A2	0	0	0	3	0	0	5	2
INV T 15	A4	0	0	0	3	60	95	4	53
INV T 15	A5	0	0	1	3	0	0	50	17
TMA block number	INVASIVE TUMOUR location	ER (Allred)	PR (Allred)	HER2 (HER2)	EGFR (HER2)	CK5/6 (%)	CK14 (%)	KI67 (%)	MCM

INV T 15	A6	0	0	2	3	0	20	30	17
INV T 15	A7	0	0	0	3	0	100	10	37
INV T 15	B1	0	0	0	0	0	0	50	17
INV T 15	B3	4	0	x	2	0	25	20	15
INV T 15	B4	6	0	1	2	0	0	20	7
INV T 15	B5	8	6	0	0	0	0	3	1
INV T 15	B6	8	7	2	0	0	0	15	5
INV T 15	B7	8	8	0	0	0	0	0	0
INV T 15	C2	8	7	1	0	0	0	15	5
INV T 15	C3	8	8	1	0	0	0	4	1
INV T 15	C4	3	0	0	1	0	0	x	x
INV T 15	C5	8	0	3	1	0	0	5	2
INV T 15	C6	4	0	0	3	0	0	20	7
INV T 15	C7	6	0	0	3	0	0	5	2
INV T 15	D1	8	8	3	0	0	0	8	3
INV T 15	D2	8	7	0	0	0	0	5	2
INV T 15	D3	7	8	2	1	0	0	4	1
INV T 15	D4	8	6	0	1	0	0	3	1
INV T 15	D6	8	6	1	1	0	0	4	1
INV T 15	D7	8	0	3	0	0	0	4	1
INV T 15	E1	8	7	0	0	0	0	0	0
INV T 15	E2	8	8	0	0	0	0	4	1
INV T 15	E3	7	0	3	0	0	0	5	x
INV T 15	E4	6	0	0	1	0	80	30	37

TMA block number	INVASIVE TUMOUR location	ER (Allred)	PR (Allred)	HER2 (HER2)	EGFR (HER2)	CK5/6 (%)	CK14 (%)	KI67 (%)	MCM
INV T 15	E5	8	8	0	1	0	0	5	2
INV T 15	E6	4	0	0	2	0	0	35	12
INV T 15	E7	0	0	0	2	0	0	20	7
INV T 15	F1	3	0	3	0	0	0	x	x
INV T 15	F2	0	3	0	2	5	4	20	10
INV T 15	F3	8	8	3	0	0	0	8	3
INV T 15	F4	8	8	0	0	0	0	5	2
INV T 15	F5	8	4	0	1	0	0	4	1
INV T 15	F7	8	0	0	0	0	0	5	2
INV T 15	G1	4	0	3	1	0	0	15	5
INV T 15	G2	8	0	3	0	0	0	x	x
INV T 15	G3	3	0	0	2	20	0	25	15
INV T 15	G4	0	0	0	2	10	0	40	17
INV T 15	G6	8	0	2	0	0	0	5	2
INV T 15	G7	4	0	0	3	0	0	12	4

2. ER and PR stained sections were scored according to the Allred method based upon the proportion and staining intensity of tumour nuclei. A score between 0 and 3 is given for intensity of staining and a score of 0-5 for % of cells stained. These two scores are combined to give a score out of eight. At the current time, an Allred score of three or more is regarded as positive in invasive breast carcinoma, with patients with ER positive invasive breast cancers eligible for anti-oestrogen therapy (32).

HER2 and EGFR were scored using the HerceptTest™ (DAKO, Ely, UK) scoring method, based upon tumour cell membrane staining. The membranous staining of tumour cells is assigned a value of 0-3 with zero being completely negative for staining, one having less than 10% of complete tumour cell membrane staining with faint or barely perceptible staining intensity. A value of two relates to greater than 10% of cells staining with moderate, complete, membrane reactivity and a score of three is assigned when greater than 10% of cells show complete tumour cell membrane staining with strong intensity. Normal breast ducts should be negative and thus act as an internal control to ensure no over-staining has occurred. Where a score 2+ for HER2 is found the result is deemed to be equivocal. For this study chromogenic in-situ hybridisation (CISH) was selected to test those cases of DCIS with HER2 scores of 2+ for genomic amplification of the HER2 gene. Fluorescence in-situ hybridisation (FISH) is regarded by some authors as being more consistent for identifying HER2 gene amplifications in FFPE tissues (185). However, concordance has been demonstrated between chromogenic in-situ hybridisation (CISH) and FISH (186) and the former does not require a fluorescence microscope or specialist filters. In this Study the "INFORM HER2 Dual ISH DNA Probe Cocktail Assay" (Ventana, Tuscan, USA) was assessed on all cases scoring 2+ using IHC. HER2 gene status is determined by enumeration of the ratio of the HER2 gene to Chromosome 17 with a ratio of HER2 to chromosome 17 of more than 2.00 or copy number of 6 or more indicating gene amplification/HER2 positivity.

Cytokeratins were deemed as staining positively where cytoplasmic and/or membranous staining was present in tumour cells and a semi-quantitative percentage score recorded for the number of tumour cells stained.

Ki67 (MIB1) staining was scored semi-quantitatively as a percentage of positive tumour cell nuclei.

Fourteen TMAs containing 20 cases of DCIS each were stained in triplicate giving a total of 840 cores for assessment for each antibody. The use of large 2.0mm cores gives a greater area than generally deemed representative in the literature, as described above. For this reason for scoring purposes only one core was required to meet the criteria of adequate for inclusion in analysis. Where two or three cores were available either an average or the maximum/highest score was recorded dependent upon the protein under evaluation (see below).

3.7.1 Analysis of differing IHC Scores.

Each antibody had a potential of three TMA cores requiring a score. This could, depending upon tissue heterogeneity, result in different scores for each TMA core per case. This posed a dilemma for analysis. As previously stated, the consensus in the literature indicates that a single core in this study would be regarded as representative but three separate scores in a single case would question this assumption. If variation between cores existed in a single lesion as scored by an individual observer then how should the final score be assigned? Should the highest or an average score of cores available be taken for each antibody?

There are guidelines for ER, PR and HER2 scoring of invasive tumours, with clinical cut-off values relating to suitability for targeted therapy. For ER and PR, the Allred method of scoring was chosen, as it is the recommended scoring system from the UK NHS Breast Screening Program for invasive breast cancer (NHSBSP guidelines

2005) for invasive cancer. For these antibodies the same rationale for scoring was applied to DCIS (Table 3).

The Allred of scoring was chosen, as it is the recommended scoring system from the UK NHS Breast Screening Program for invasive breast cancer (NHSBSP guidelines 2005) for invasive cancer. For these antibodies the same rationale for scoring was applied to DCIS).

Similarly, as the scoring for EGFR mirrored HER2 in terms of pattern, the same cut-off principles were applied. Cut-off values for the basal cytokeratins and Ki67 were harder to determine, with significant variation in the literature (see discussion). For these an average score of all representative cores was taken.

Currently there are limited data on the value of TMAs for setting quantitative relationships or cut-off values to clinical samples and this invariably plays a role in comparison of different series. The International Ki67 in Breast Cancer working group (187) suggest that TMAs should not be used to predict cut-offs, however, the use of 2mm cores as opposed to 0.6mm may somewhat ameliorate any discrepancy between whole sections and TMAs. In this series the median value was chosen as a cut-off due to the lack on consensus for a clinical cut-off value for Ki67 in cases of DCIS.

3.7.2 Inter Observer Variability in Scoring Immunohistochemistry

All cores were independently scored by the two reviewers. Comparison of results between the two reviewers was performed to determine reproducibility. The degree of variance allowed and the best scoring method needed to be defined prior to

screening to eliminate any possible bias; the allowances for inter observer variation are given in 4. Scores that differed between individuals greater than the accepted variance underwent joint review by both scorers with a consensus being reached.

Score for proportion	Score for intensity
0 = no staining	0 = no staining
1 = <1% nuclei staining	1 = weak staining
2 = 1-10% nuclei staining	2 = moderate staining
3 = 11-33% nuclei staining	3 = strong staining
4 = 34-66% nuclei staining	
5 = 67-100% nuclei staining	

Table 3: Allred Scoring Method

3.8 Statistical Analysis of Immunohistochemical Scoring

Statistical analyses were performed using the software package StatView (SAS Institute Inc. San Francisco, CA). Frequency distributions and non parametric univariate analysis was undertaken to compare the expression the biomarkers with each other and with the pathological criteria utilising the Chi-square (χ^2) and Fisher's exact tests as appropriate. Fisher's exact test is suitable for defining a p value where two categorical variables have small values (usually less than 5). Chi- square is used to test for associations between larger values. Frequency distributions and univariate analyses tables are listed in Appendix 5.

3.9 Immunohistochemistry Results

3.9.1 Immunohistochemical Scoring

Immunohistochemical results for staining of DCIS are given below (Table 5) with the number of positive and negative stained cases according to the scoring criteria above. Staining frequencies for each antibody are listed in Table 5. Representative images for IHC staining of pure DCIS & CISH for HER2 are seen in Figure 23 to Figure 29 (section 3.9.2) Sections 3.9.3 to 3.9.9 give representative examples of paired univariate analysis for IHC markers and observed frequencies. Appendix 5 lists all other sets.

Antibodies	Scoring Method, recorded score	Allowed variance between observers
ER and PR	Allred (0-8). The average of all cores is taken.	Scores must be within 1 of each other i.e. a score of 7 on one core and 6 or 8 on the others is acceptable.
HER2 and EGFR	HER2 score, 0-3. Highest value is regarded as definitive.	No variance accepted
CKs (5, 5/6, 14)	%, an average of scores from the cores is regarded as final result.	Variance of 5% is acceptable
Ki67	%, an average of scores from the cores is regarded as final score.	Variance of 5% is acceptable

Table 4: IHC methods of scoring, of handling differing scores from cores and defined acceptable variation for DCIS immunostaining.

Antibody	Total Number of Cases Stained	Number of Positive Cases (%)	Number of Negative Cases (%)
ER	224	165 (74)	59 (26)
PR	236	113 (48)	123 (52)
HER2	248	57 (23)	191 (77)
EGFR	227	11 (5)	213 (95)
CK5	227	32 (14)	195 (86)
CK5/6	220	69 (31)	151 (69)
CK14	231	38 (16)	193 (84)
Ki67	237	136 (57)	101 (43)
MCM2	237	130 (55)	107 (45)

Table 5: Staining frequencies for immunohistochemical antibody staining.

3.9.2 Immunohistochemical Images

Figure 23: ER staining of DCIS in 2mm TMA core (X40, X200)

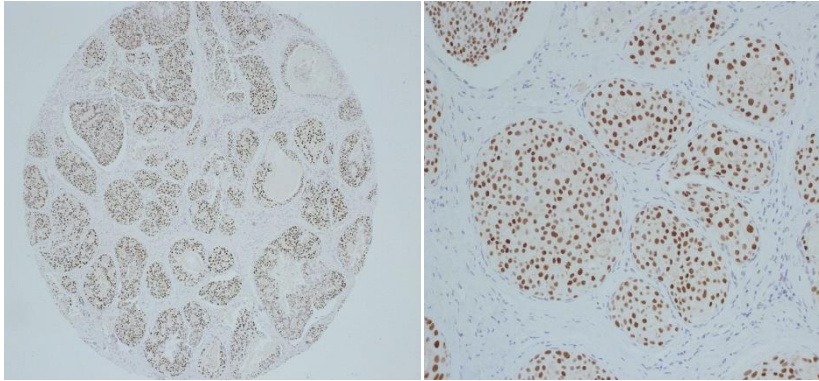


Figure 24: PR staining of DCIS in 2mm TMA core (X40, X200)

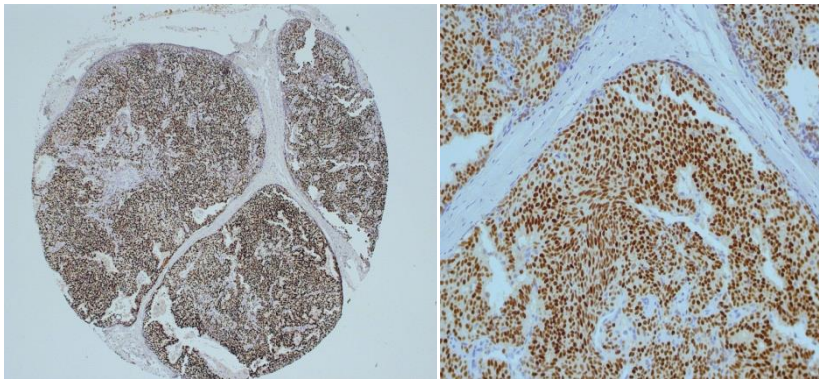


Figure 25: HER2 staining of DCIS in 2mm TMA core (X100, X200)

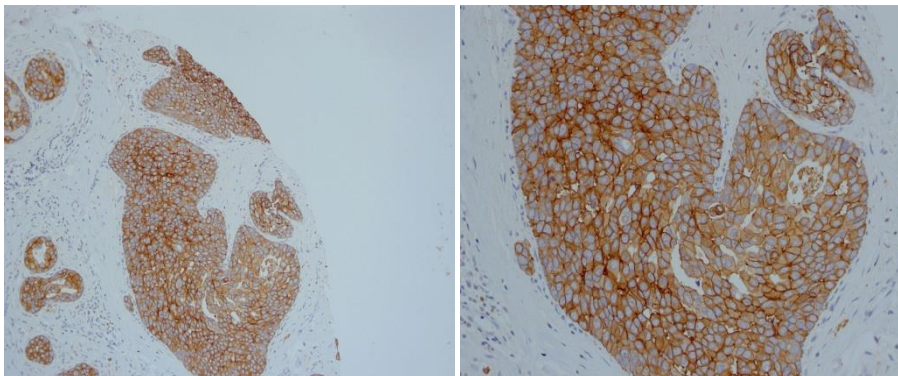


Figure 26: EGFR staining of DCIS in 2mm TMA core (X40, X100, X200)

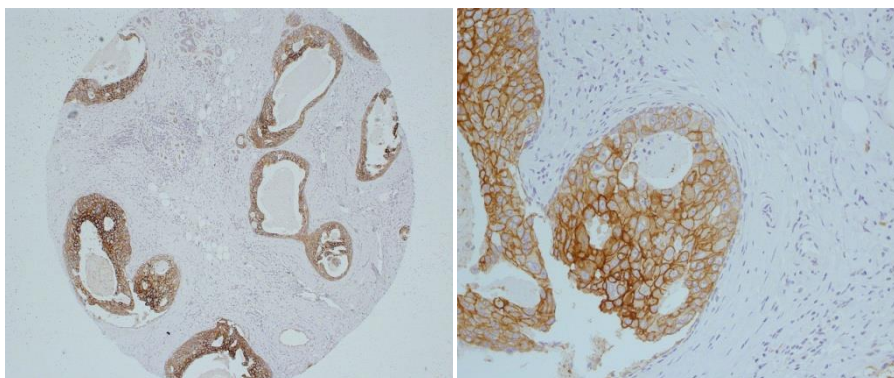


Figure 27: CK14 staining of DCIS in 2mm TMA core (X40, X200).

[Peripheral myoepithelial cell staining was not included in assessment of the DCIS]

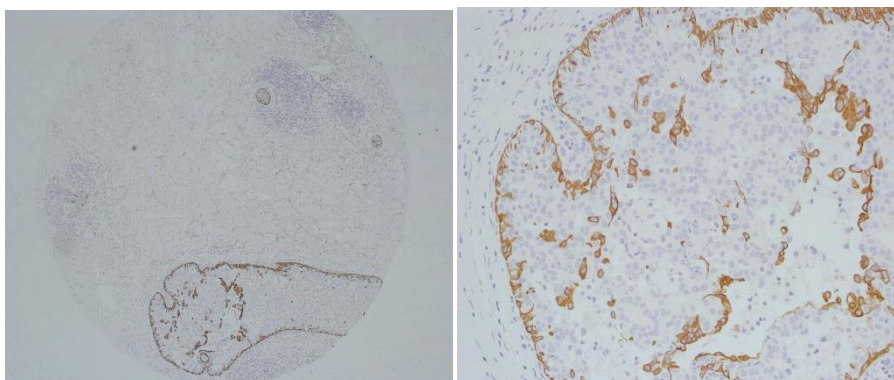


Figure 28: Ki67 staining of DCIS in 2mm TMA core (X40, X200)

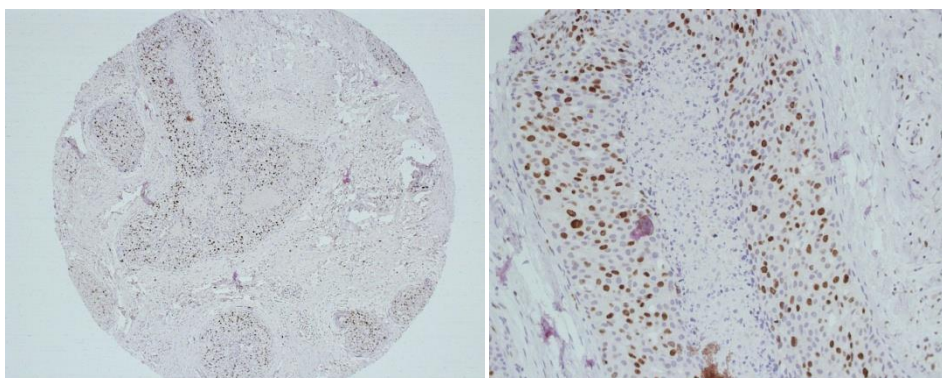
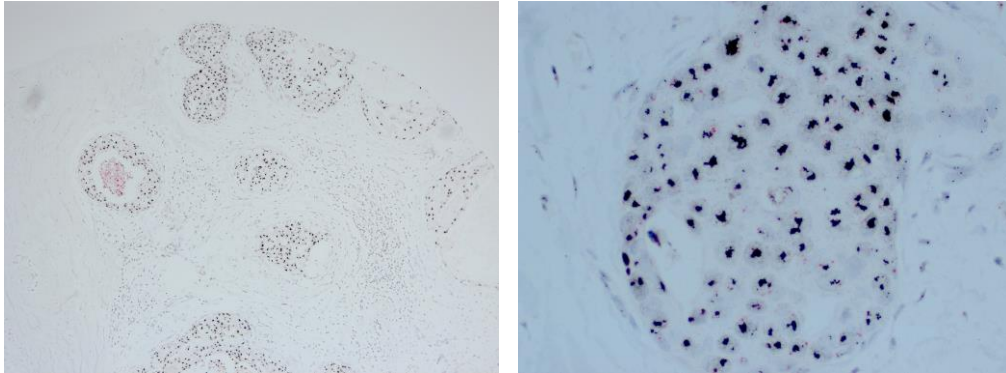


Figure 29: CISH for HER2 showing amplification of the *HER2* Gene in DCIS, 2mm Core (x40, X200)



3.9.3 Non Parametric Tests- Univariate Paired analysis:

Positive univariate paired analyses tables and p-values for ER, PR, HER2, EGFR, Ki67, MCM2, CK5, CK5/6, CK14 and histological parameters are given below (Table 6).

Antibody	HER2	ER	PR	Ki67	Grade	Inflam- mation	Nec- rosis	MCM2	Archit- ecture	CK5/6	CK14	CK5	EGFR
HER2													
ER	<0.000 1												
PR	<0.000 1	<0.000 1											
Ki67	<0.000 1	<0.000 1	<0.000 1										
Grade	0.0001	<0.000 1	0.0006	0.0009									
Inflammation	<0.000 1	<0.000 1	<0.000 1	<0.000 1	<0.000 1								
Necrosis	<0.000 1	<0.000 1	0.0013	<0.000 1	<0.000 1	<0.000 1							
MCM2	<0.000 1	0.0032	0.0033	<0.000 1	<0.000 1	0.0003	<0.000 1						
Architecture	0.0003	<0.000 1	0.0011	0.0073	<0.000 1	0.0025	0.0008	0.0098					
CK5/6	0.0041	0.0298	0.0542	0.2279	0.1044	0.0025	0.0515	0.3772	0.006 6				
CK14	0.0037	0.0018	0.2791	0.29	0.115	0.2056	0.042	0.215	0.135 5	<0.000 1			
CK5	0.0443	0.0033	0.2385	0.3305	0.1364	0.0422	0.804	0.4392	0.201 1	<0.000 1	0.0003		
EGFR	0.7324	<0.000 1	0.0008	0.0294	0.0865	0.0126	0.0816	0.0254	0.152 7	>0.999 9	>0.999 9	0.2139	
Ipsilateral Recurrence	>0.999 9	0.2912	0.3666	0.0812	0.3851	0.6047	0.4063	>0.999 9	0.483 4	>0.999 9	0.7327	0.6594	>0.999 9
All Recurrence	0.1952	0.784	>0.999 9	0.4856	0.8452	0.8269	0.8352	0.4938	0.926 6	>0.999 9	0.4113	>0.999 9	>0.999 9

Table 6: Univariate analysis p-values for immunohistochemical and histological parameters.

3.9.4 Univariate Analysis of Hormone Receptor Immunohistochemistry

Summary Table for FINAL ER STATUS, FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	52
DF	1
Chi Square	57.061
Chi Square P-Value	<.0001
G-Squared	52.622
G-Squared P-Value	<.0001
Contingency Coef.	.451
Phi	.506
Cty. Cor. Chi Square	54.443
Cty. Cor. P-Value	<.0001
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL ER STATUS, FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	23	36	59
1	145	19	164
Totals	168	55	223

Summary Table for FINAL EGFR STATUS, FINAL ER STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	60
DF	1
Chi Square	25.989
Chi Square P-Value	<.0001
G-Squared	22.528
G-Squared P-Value	<.0001
Contingency Coef.	.328
Phi	.348
Cty. Cor. Chi Square	22.514
Cty. Cor. P-Value	<.0001
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL EGFR STATUS, FINAL ER STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	45	159	204
1	10	1	11
Totals	55	160	215

Table 7: Univariate paired analysis for oestrogen receptor immunohistochemistry

Summary Table for FINAL PR STATUS, FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	45
DF	1
Chi Square	40.858
Chi Square P-Value	<.0001
G-Squared	45.748
G-Squared P-Value	<.0001
Contingency Coef.	.388
Phi	.421
Cty. Cor. Chi Square	38.915
Cty. Cor. P-Value	<.0001
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL PR STATUS, FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	70	50	120
1	104	6	110
Totals	174	56	230

Summary Table for FINAL EGFR STATUS, FINAL PR STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	57
DF	1
Chi Square	10.764
Chi Square P-Value	.0010
G-Squared	.
G-Squared P-Value	.
Contingency Coef.	.217
Phi	.222
Cty. Cor. Chi Square	8.829
Cty. Cor. P-Value	.0030
Fisher's Exact P-Value	.0008

Observed Frequencies for FINAL EGFR STATUS, FINAL PR STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	102	105	207
1	11	0	11
Totals	113	105	218

Table 8: Univariate paired analysis for progesterone receptor immunohistochemistry. The majority of DCIS cases were positive for hormone receptors (ER n=165, PR n=144- Table 7Table 8). For ER status there was an inverse correlation ($p=0.001$) with both EGFR and HER2 (Table 9); ER positive DCIS was associated with a negative EGFR and/or HER2 status. All PR positive cases were also ER positive

(n=102 Table 10) There was an inverse correlation for CK5 and CK14 and ER with 90% (n=147) and 88% (n=142) of ER positive cases being CK5 and CK14 negative respectively (Table 12). Few cases were negative for both ER and proliferation markers (ER-ve + Ki67-ve n=8 Table 13 and ER-ve MCM –ve n=15

Table 14) All low grade DCIS (n=9) was ER positive (6% of all graded cases) and 82% (n=46) of intermediate grade DCIS (n=50) was ER positive (34% of all graded cases; Table 15). High grade DCIS (n=90) made up 60% of cases and 61% (n=55) of these lesions were ER positive (p<0.0001). ER positive cases were more likely to have cribriform (n=55) or solid (n=38) architecture (Table 16). ER status did not predict the likelihood of recurrence of either DCIS or invasive disease (p=0.784; Table 20).

	ER -ve	ER +ve	Total
HER2 -ve	23	145	168
HER2 +ve	36	19	55
Total	59	164	223

Table 9: Univariate analyses for ER and HER2 IHC: Fishers exact test p value < 0.0001

	ER -ve	ER +ve	Total
PR -ve	56	55	111
PR +ve	0	102	102
Total	56	157	213

Table 10: Univariate analysis for ER and PR IHC; Fishers exact test p value <0.0001

	ER -ve	ER +ve	Total
CK5 -ve	41	147	188
CK5 +ve	15	16	31
Total	56	163	219

Table 11: Univariate analyses for ER and CK5 IHC; Fishers exact test p value = 0.0033.

	ER -ve	ER +ve	Total
CK14 -ve	40	142	182
CK14 +ve	18	19	37
Total	58	161	219

Table 12: Univariate analyses for ER and CK14 IHC; Fishers exact test p value = 0.0018.

	ER -ve	ER +ve	Total
Ki67 -ve	8	78	86
Ki67+ve	51	79	130
Total	59	157	216

Table 13: Univariate analyses for ER and Ki67 (>5%) IHC; Fishers exact test p value <0.0001

	ER -ve	ER +ve	Total
MCM2 -ve	15	77	92
MCM2+ve	44	83	127
Total	59	160	219

Table 14: Univariate analysis for ER and MCM2; (>5%) IHC Fishers exact test p value = 0.0032

	ER -ve	ER +ve	Total
HG	35	55	90
IG	4	46	50
LG	0	9	9
Total	39	110	149

Table 15: Univariate analyses for ER and Histological Grade; ER; Chi Square p value <0.0001

	ER -ve	ER +ve	Total
Crib	5	55	60
Micropap	7	6	13
Pap	1	11	12
Solid	26	38	64
Total	39	110	149

Table 16: Univariate analyses for ER IHC and Histological Architecture; Chi square p Value < 0.0001

3.9.5 Univariate Analysis for HER2 Immunohistochemistry

Summary Table for FINAL HER2 STATUS, FINAL EGFR STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	50
DF	1
Chi Square	.050
Chi Square P-Value	.8229
G-Squared	.049
G-Squared P-Value	.8249
Contingency Coef.	.015
Phi	.015
Cty. Cor. Chi Square	0.000
Cty. Cor. P-Value	>.9999
Fisher's Exact P-Value	.7324

Observed Frequencies for FINAL HER2 STATUS, FINAL EGFR STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	162	8	170
1	52	3	55
Totals	214	11	225

Summary Table for FINAL HER2 STATUS, FINAL ER STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	52
DF	1
Chi Square	57.061
Chi Square P-Value	<.0001
G-Squared	52.622
G-Squared P-Value	<.0001
Contingency Coef.	.451
Phi	.506
Cty. Cor. Chi Square	54.443
Cty. Cor. P-Value	<.0001
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL HER2 STATUS, FINAL ER STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	23	145	168
1	36	19	55
Totals	59	164	223

Table 17: Univariate paired analysis for HER2 immunohistochemistry.

A total of 248 cases had HER2 IHC results (Table 17), of these 23% (n=57) were HER2 positive. This included 5% (n=13) of the 2+ cases on IHC (n=47) which were subsequently shown to show *HER2* gene amplification by CISH. Of the HER2 positive cases 9% (n=19) were also ER positive. 24% (n=39) of patients for whom clinical data was available (n=161) had HER2 positive DCIS. Of these HER2 positive cases 28% (n=11) were also ER positive ($p<0.0001$) 13% (n=5) was of intermediate grade ($p=0.0001$). Of note, 4 of the 5 cases of intermediate grade DCIS which were HER2 positive were also ER positive (80%). There were no low grade HER2 positive DCIS cases. A significant p value was given between HER2 positivity and architectural pattern ($p=0.0003$), the presence of necrosis ($p=0.0001$) and chronic inflammation ($p<0.001$). None of the 11 HER2/ER positive cases had an invasive recurrence, 45% (n=5) underwent mastectomy as a result of the DCIS diagnosis and 3 had recurrent DCIS in the same year. Three cases of DCIS were both HER2 and EGFR positive 7%, all had a simple mastectomy.

3.9.6 Univariate Analysis for Proliferation Marker Immunohistochemistry

Summary Table for FINAL KI67 (AT MEDIAN, 5%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	43
DF	1
Chi Square	15.451
Chi Square P-Value	<.0001
G-Squared	16.560
G-Squared P-Value	<.0001
Contingency Coef.	.250
Phi	.258
Cty. Cor. Chi Square	14.255
Cty. Cor. P-Value	.0002
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL KI67 (AT MEDIAN, 5%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	87	11	98
1	89	45	134
Totals	176	56	232

Summary Table for FINAL EGFR STATUS, FINAL KI67 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	53
DF	1
Chi Square	4.869
Chi Square P-Value	.0273
G-Squared	5.872
G-Squared P-Value	.0154
Contingency Coef.	.146
Phi	.148
Cty. Cor. Chi Square	3.580
Cty. Cor. P-Value	.0585
Fisher's Exact P-Value	.0294

Observed Frequencies for FINAL EGFR STATUS, FINAL KI67 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	90	121	211
1	1	10	11
Totals	91	131	222

Summary Table for FINAL MCM2 (AT MEDIAN, 27.5%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	39
DF	1
Chi Square	19.314
Chi Square P-Value	<.0001
G-Squared	20.682
G-Squared P-Value	<.0001
Contingency Coef.	.275
Phi	.286
Cty. Cor. Chi Square	17.992
Cty. Cor. P-Value	<.0001
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL MCM2 (AT MEDIAN, 27.5%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	94	11	105
1	85	46	131
Totals	179	57	236

Summary Table for FINAL EGFR STATUS, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	50
DF	1
Chi Square	5.588
Chi Square P-Value	.0181
G-Squared	6.671
G-Squared P-Value	.0098
Contingency Coef.	.156
Phi	.158
Cty. Cor. Chi Square	4.217
Cty. Cor. P-Value	.0400
Fisher's Exact P-Value	.0254

Observed Frequencies for FINAL EGFR STATUS, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	97	117	214
1	1	10	11
Totals	98	127	225

Table 18: Univariate Analysis of Proliferation Marker Immunohistochemistry.

The proliferation markers examined (Table 18), Ki67 and MCM2, were both scored as positive when >5% of DCIS cells showed nuclear positivity, which corresponded to the median value for Ki67. 91% (n=78) of Ki67 negative cases were positive for ER. Only 4% (n=8) of cases were negative for both Ki67 and ER ($p<0.0001$).

The association between PR and Ki67 was similar to ER and Ki67; 16% (n=36) cases were negative for both PR and Ki67 ($p<0.0001$). If one of these markers was positive the other was most likely to be negative. MCM2 and ER showed a similar profile to Ki67 and ER, with 7% (n=15) of cases negative for both markers and 75% (n=44) of ER negative cases being MCM2 positive ($p=0.0032$). Similar to Ki67 and PR, 18% (n=42) of DCIS cases were negative for both PR and MCM2, whilst cases tended to be positive for one of these markers and negative for the other ($p=0.0033$).

3.9.7 Univariate Analysis for Cytokeratin Immunohistochemistry

Summary Table for FINAL CK5 (AT 1%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	49
DF	1
Chi Square	5.022
Chi Square P-Value	.0250
G-Squared	4.594
G-Squared P-Value	.0321
Contingency Coef.	.147
Phi	.149
Cty. Cor. Chi Square	4.081
Cty. Cor. P-Value	.0434
Fisher's Exact P-Value	.0443

Observed Frequencies for FINAL CK5 (AT 1%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	151	43	194
1	19	13	32
Totals	170	56	226

Summary Table for FINAL EGFR STATUS, FINAL CK5 (AT 1%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	60
DF	1
Chi Square	1.405
Chi Square P-Value	.2360
G-Squared	1.190
G-Squared P-Value	.2754
Contingency Coef.	.081
Phi	.081
Cty. Cor. Chi Square	.566
Cty. Cor. P-Value	.4518
Fisher's Exact P-Value	.2139

Observed Frequencies for FINAL EGFR STATUS, FINAL CK5 (AT 1%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	175	29	204
1	8	3	11
Totals	183	32	215

Summary Table for FINAL CK5/6 (AT 1%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	58
DF	1
Chi Square	8.863
Chi Square P-Value	.0029
G-Squared	8.507
G-Squared P-Value	.0035
Contingency Coef.	.198
Phi	.202
Cty. Cor. Chi Square	7.892
Cty. Cor. P-Value	.0050
Fisher's Exact P-Value	.0041

Observed Frequencies for FINAL CK5/6 (AT 1%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	120	28	148
1	43	26	69
Totals	163	54	217

Summary Table for FINAL EGFR STATUS, FINAL CK5/6 (AT 1%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	62
DF	1
Chi Square	.139
Chi Square P-Value	.7094
G-Squared	.143
G-Squared P-Value	.7053
Contingency Coef.	.026
Phi	.026
Cty. Cor. Chi Square	.002
Cty. Cor. P-Value	.9653
Fisher's Exact P-Value	>.9999

Observed Frequencies for FINAL EGFR STATUS, FINAL CK5/6 (AT 1%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	136	66	202
1	8	3	11
Totals	144	69	213

Summary Table for FINAL CK14 (AT 1%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	47
DF	1
Chi Square	9.474
Chi Square P-Value	.0021
G-Squared	8.599
G-Squared P-Value	.0034
Contingency Coef.	.200
Phi	.204
Cty. Cor. Chi Square	8.253
Cty. Cor. P-Value	.0041
Fisher's Exact P-Value	.0037

Observed Frequencies for FINAL CK14 (AT 1%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	150	40	190
1	21	17	38
Totals	171	57	228

Summary Table for FINAL EGFR STATUS, FINAL CK14 (AT 1%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	55
DF	1
Chi Square	.007
Chi Square P-Value	.9348
G-Squared	.007
G-Squared P-Value	.9352
Contingency Coef.	.006
Phi	.006
Cty. Cor. Chi Square	0.000
Cty. Cor. P-Value	>.9999
Fisher's Exact P-Value	>.9999

Observed Frequencies for FINAL EGFR STATUS, FINAL CK14 (AT 1%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	173	36	209
1	9	2	11
Totals	182	38	220

Table 19: Paired univariate analysis for CK immunohistochemistry.

The frequency of expression of CK5 and CK5/6 was distinctly different; 31% of cases were positive for CK5/6 (n=69) compared to only 14% for CK5 (n=32)(Table 19). Few cases showed EGFR positivity (n=11, 5%); 3 EGFR positive cases were also either CK5/6 or CK5 positive and 2 cases exhibited both CK14 and EGFR positivity.

Of the basal cytokeratin markers, CK14 ($p=0.0018$) and CK5 ($p=0.0033$) demonstrated an inverse relationship with ER but no significant association between CK5/6 ($p=0.298$) and ER was seen. Over 90% of EGFR positive cases were ER negative and over 99% of ER positive cases were EGFR negative ($p<0.0001$). All but one case of EGFR positive DCIS was positive for both proliferation markers Ki67 ($p=0.0294$) and MCM2 ($p=0.0254$) EGFR positive DCIS was of higher cytonuclear grade and was more likely to have solid architecture with comedo-type necrosis and associated chronic inflammation present.

3.9.8 Univariate Analysis and Recurrence of DCIS and/or Invasive Breast Disease

Univariate paired analyses were carried out for all antibodies to determine any significant correlation with recurrence (Table 20).

Summary Table for REC Y/N, FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	121
DF	1
Chi Square	2.156
Chi Square P-Value	.1420
G-Squared	2.426
G-Squared P-Value	.1194
Contingency Coef.	.118
Phi	.118
Cty. Cor. Chi Square	1.461
Cty. Cor. P-Value	.2268
Fisher's Exact P-Value	.1952

Observed Frequencies for REC Y/N, FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
N	95	36	131
Y	20	3	23
Totals	115	39	154

Summary Table for FINAL EGFR STATUS, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	130
DF	1
Chi Square	.084
Chi Square P-Value	.7723
G-Squared	.080
G-Squared P-Value	.7773
Contingency Coef.	.024
Phi	.024
Cty. Cor. Chi Square	0.000
Cty. Cor. P-Value	>.9999
Fisher's Exact P-Value	.6738

Observed Frequencies for FINAL EGFR STATUS, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	114	20	134
1	9	2	11
Totals	123	22	145

Summary Table for FINAL ER STATUS, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	135
DF	1
Chi Square	.152
Chi Square P-Value	.6971
G-Squared	.148
G-Squared P-Value	.7002
Contingency Coef.	.033
Phi	.033
Cty. Cor. Chi Square	.013
Cty. Cor. P-Value	.9092
Fisher's Exact P-Value	.7840

Observed Frequencies for FINAL ER STATUS, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	33	6	39
1	88	13	101
Totals	121	19	140

Summary Table for FINAL PR STATUS, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	125
DF	1
Chi Square	.004
Chi Square P-Value	.9489
G-Squared	.004
G-Squared P-Value	.9489
Contingency Coef.	.005
Phi	.005
Cty. Cor. Chi Square	0.000
Cty. Cor. P-Value	>.9999
Fisher's Exact P-Value	>.9999

Observed Frequencies for FINAL PR STATUS, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	64	10	74
1	66	10	76
Totals	130	20	150

Summary Table for FINAL CK5 (AT 1%), REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	130
DF	1
Chi Square	.005
Chi Square P-Value	.9436
G-Squared	.005
G-Squared P-Value	.9438
Contingency Coef.	.006
Phi	.006
Cty. Cor. Chi Square	0.000
Cty. Cor. P-Value	>.9999
Fisher's Exact P-Value	>.9999

Observed Frequencies for FINAL CK5 (AT 1%), REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	107	17	124
1	18	3	21
Totals	125	20	145

Summary Table for FINAL CK5/6 (AT 1%), REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	133
DF	1
Chi Square	.061
Chi Square P-Value	.8050
G-Squared	.062
G-Squared P-Value	.8039
Contingency Coef.	.021
Phi	.021
Cty. Cor. Chi Square	0.000
Cty. Cor. P-Value	>.9999
Fisher's Exact P-Value	>.9999

Observed Frequencies for FINAL CK5/6 (AT 1%), REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	82	14	96
1	40	6	46
Totals	122	20	142

Summary Table for FINAL CK14 (AT 1%), REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	126
DF	1
Chi Square	.941
Chi Square P-Value	.3321
G-Squared	1.028
G-Squared P-Value	.3106
Contingency Coef.	.079
Phi	.079
Cty. Cor. Chi Square	.477
Cty. Cor. P-Value	.4898
Fisher's Exact P-Value	.4113

Observed Frequencies for FINAL CK14 (AT 1%), REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	98	19	117
1	29	3	32
Totals	127	22	149

Summary Table for FINAL KI67 (AT MEDIAN, 5%, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	128
DF	1
Chi Square	.632
Chi Square P-Value	.4267
G-Squared	.646
G-Squared P-Value	.4214
Contingency Coef.	.065
Phi	.066
Cty. Cor. Chi Square	.314
Cty. Cor. P-Value	.5755
Fisher's Exact P-Value	.4856

Observed Frequencies for FINAL KI67 (AT MEDIAN, 5%, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	51	7	58
1	74	15	89
Totals	125	22	147

Summary Table for FINAL MCM2 (AT MEDIAN, 27.5%, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	124
DF	1
Chi Square	.508
Chi Square P-Value	.4761
G-Squared	.504
G-Squared P-Value	.4778
Contingency Coef.	.058
Phi	.058
Cty. Cor. Chi Square	.232
Cty. Cor. P-Value	.6303
Fisher's Exact P-Value	.4938

Observed Frequencies for FINAL MCM2 (AT MEDIAN, 27.5%), REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	54	11	65
1	75	11	86
Totals	129	22	151

Summary Table for GRADE, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

Num. Missing	118
DF	2
Chi Square	.336
Chi Square P-Value	.8452
G-Squared	.370
G-Squared P-Value	.8310
Contingency Coef.	.046
Cramer's V	.046

Observed Frequencies for GRADE, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
HG	78	13	91
IG	47	7	54
LG	11	1	12
Totals	136	21	157

Summary Table for ARCH, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

Num. Missing	118
DF	3
Chi Square	.465
Chi Square P-Value	.9266
G-Squared	.439
G-Squared P-Value	.9322
Contingency Coef.	.054
Cramer's V	.054

Observed Frequencies for ARCH, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
CRIB	56	8	64
MICROPAP	14	3	17
PAP	10	2	12
SOLID	56	8	64
Totals	136	21	157

Summary Table for NECROSIS, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	118
DF	3
Chi Square	.860
Chi Square P-Value	.8352
G-Squared	.914
G-Squared P-Value	.8219
Contingency Coef.	.074
Cramer's V	.074

Observed Frequencies for NECROSIS, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
MARKED	35	6	41
MILD	39	4	43
MOD	29	5	34
NONE	33	6	39
Totals	136	21	157

Summary Table for CI, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	120
DF	3
Chi Square	.894
Chi Square P-Value	.8269
G-Squared	.947
G-Squared P-Value	.8142
Contingency Coef.	.076
Cramer's V	.076

Observed Frequencies for CI, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
MARKED	32	3	35
MILD	47	7	54
MOD	29	5	34
NONE	27	5	32
Totals	135	20	155

Summary Table for REC Y/N, SIZE RECODED AT MEDIAN (15)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	125
DF	1
Chi Square	.853
Chi Square P-Value	.3556
G-Squared	.854
G-Squared P-Value	.3553
Contingency Coef.	.075
Phi	.075
Cty. Cor. Chi Square	.460
Cty. Cor. P-Value	.4977
Fisher's Exact P-Value	.4625

Observed Frequencies for REC Y/N, SIZE RECODED AT MEDIAN (15)

Row exclusion: DCIS STATVIEW DATASET

	LESS THAN MEDIAN	MEDIAN (15) OR MORE	Totals
N	61	70	131
Y	11	8	19
Totals	72	78	150

Summary Table for AGE RECODED AT 50, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	121
DF	1
Chi Square	.414
Chi Square P-Value	.5201
G-Squared	.402
G-Squared P-Value	.5261
Contingency Coef.	.052
Phi	.052
Cty. Cor. Chi Square	.156
Cty. Cor. P-Value	.6924
Fisher's Exact P-Value	.6208

Observed Frequencies for AGE RECODED AT 50, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
AGE 50 OR LESS	35	8	43
MORE THAN 50	95	16	111
Totals	130	24	154

Summary Table for LUMINAL A, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	138
DF	1
Chi Square	.052
Chi Square P-Value	.8204
G-Squared	.052
G-Squared P-Value	.8204
Contingency Coef.	.019
Phi	.019
Cty. Cor. Chi Square	0.000
Cty. Cor. P-Value	>.9999
Fisher's Exact P-Value	>.9999

Observed Frequencies for LUMINAL A, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	60	8	68
1	60	9	69
Totals	120	17	137

Summary Table for LUMINAL B, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	131
DF	1
Chi Square	1.042
Chi Square P-Value	.3073
G-Squared	1.177
G-Squared P-Value	.2780
Contingency Coef.	.085
Phi	.085
Cty. Cor. Chi Square	.500
Cty. Cor. P-Value	.4797
Fisher's Exact P-Value	.5290

Observed Frequencies for LUMINAL B, REC Y/N

Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	99	16	115
1	27	2	29
Totals	126	18	144

Summary Table for HER2 GROUP, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

Num. Missing	135
DF	1
Chi Square	.173
Chi Square P-Value	.6778
G-Squared	.180
G-Squared P-Value	.6713
Contingency Coef.	.035
Phi	.035
Cty. Cor. Chi Square	.011
Cty. Cor. P-Value	.9182
Fisher's Exact P-Value	>.9999

Observed Frequencies for HER2 GROUP, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	97	16	113
1	24	3	27
Totals	121	19	140

Summary Table for LUMINAL-HER2, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

Num. Missing	135
DF	1
Chi Square	1.875
Chi Square P-Value	.1710
G-Squared	.
G-Squared P-Value	.
Contingency Coef.	.115
Phi	.116
Cty. Cor. Chi Square	.829
Cty. Cor. P-Value	.3625
Fisher's Exact P-Value	.3605

Observed Frequencies for LUMINAL-HER2, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	110	19	129
1	11	0	11
Totals	121	19	140

Summary Table for BASAL-LIKE GROUP, REC Y/N**Row exclusion: DCIS STATVIEW DATASET**

Num. Missing	136
DF	1
Chi Square	.181
Chi Square P-Value	.6709
G-Squared	.187
G-Squared P-Value	.6658
Contingency Coef.	.036
Phi	.036
Cty. Cor. Chi Square	.020
Cty. Cor. P-Value	.8863
Fisher's Exact P-Value	.7830

Observed Frequencies for BASAL-LIKE GROUP, REC Y/N**Row exclusion: DCIS STATVIEW DATASET**

	N	Y	Totals
0	90	16	106
1	29	4	33
Totals	119	20	139

Summary Table for LUMINAL B1, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

Num. Missing	141
DF	1
Chi Square	.085
Chi Square P-Value	.7705
G-Squared	.084
G-Squared P-Value	.7716
Contingency Coef.	.025
Phi	.025
Cty. Cor. Chi Square	.001
Cty. Cor. P-Value	.9780
Fisher's Exact P-Value	.7955

Observed Frequencies for LUMINAL B1, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	75	11	86
1	41	7	48
Totals	116	18	134

Summary Table for LUMINAL A1, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

Num. Missing	141
DF	1
Chi Square	.486
Chi Square P-Value	.4856
G-Squared	.502
G-Squared P-Value	.4787
Contingency Coef.	.060
Phi	.060
Cty. Cor. Chi Square	.186
Cty. Cor. P-Value	.6659
Fisher's Exact P-Value	.6004

Observed Frequencies for LUMINAL A1, REC Y/N
Row exclusion: DCIS STATVIEW DATASET

	N	Y	Totals
0	74	13	87
1	42	5	47
Totals	116	18	134

Table 20: Recurrence data for IHC Markers and Luminal Groups

3.9.9 Histological Grade in DCIS

Immunohistochemical expression of all markers in this study was compared to histological grade to determine any correlation between grade and specific markers. Table 21 gives the number of positive and negative cases expressing IHC markers in relation to cytonuclear grade.

	High Grade (n)	High Grade (% of total)	Intermediate Grade (n)	Intermediate Grade (% of total)	Low Grade (n)	Low Grade (% of total)	Total (n)
ER -ve	35	90	4	10	0	0	39
ER +ve	55	50	46	42	9	8	110
Total	90		50		9		149
PR -ve	54	70	23	30	0	0	77
PR +ve	37	46	33	41	10	13	80
Total	91		56		10		157
EGFR -ve	81	57	53	37	8	6	142
EGFR +ve	10	91	1	9	0	0	11
Total	91		54		8		153
HER2 -ve	60	49	51	41	12	10	123
HER2 +ve	33	87	5	13	0	0	38
Total	93		56		12		161
Ki67 low	26	42	28	45	8	13	62
Ki67 high	65	71	24	26	3	3	92
Total	91		52		11		154
MCM2 low	28	39	36	51	7	10	71
MCM2 high	66	76	17	20	4	5	87
Total	94		53		11		158
CK5 -ve	73	57	46	36	10	8	129
CK5 +ve	17	77	5	23	0	0	22
Total	90		51		10		151
CK5/6 -ve	56	54	42	41	5	5	103
CK5/6 +ve	34	72	11	23	2	4	47
Total	90		53		7		150
CK14 -ve	69	55	49	39	7	6	125
CK14 +ve	23	74	6	19	2	6	31
Total	92		55		9		156

Table 21: Biomarker expression by cytonuclear grade of DCIS.

Paired univariate analysis between the immunohistochemical markers and histological parameters such as cytonuclear grade, comedo type necrosis, chronic inflammation was undertaken to determine any correlations (Table 22).

Summary Table for GRADE, FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	114
DF	2
Chi Square	17.675
Chi Square P-Value	.0001
G-Squared	•
G-Squared P-Value	•
Contingency Coef.	.315
Cramer's V	.331

Observed Frequencies for GRADE, FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
HG	60	33	93
IG	51	5	56
LG	12	0	12
Totals	123	38	161

Summary Table for FINAL EGFR STATUS, GRADE

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	122
DF	2
Chi Square	4.894
Chi Square P-Value	.0865
G-Squared	•
G-Squared P-Value	•
Contingency Coef.	.176
Cramer's V	.179

Observed Frequencies for FINAL EGFR STATUS, GRADE

Row exclusion: DCIS STATVIEW DATASET

	HG	IG	LG	Totals
0	81	53	8	142
1	10	1	0	11
Totals	91	54	8	153

Summary Table for FINAL ER STATUS, GRADE

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	126
DF	2
Chi Square	19.267
Chi Square P-Value	<.0001
G-Squared	•
G-Squared P-Value	•
Contingency Coef.	.338
Cramer's V	.360

Observed Frequencies for FINAL ER STATUS, GRADE

Row exclusion: DCIS STATVIEW DATASET

	HG	IG	LG	Totals
0	35	4	0	39
1	55	46	9	110
Totals	90	50	9	149

Summary Table for FINAL PR STATUS, GRADE

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	118
DF	2
Chi Square	14.910
Chi Square P-Value	.0006
G-Squared	•
G-Squared P-Value	•
Contingency Coef.	.294
Cramer's V	.308

Observed Frequencies for FINAL PR STATUS, GRADE

Row exclusion: DCIS STATVIEW DATASET

	HG	IG	LG	Totals
0	54	23	0	77
1	37	33	10	80
Totals	91	56	10	157

Summary Table for FINAL CK5 (AT 1%), GRADE
Row exclusion: DCIS STATVIEW DATASET

Num. Missing	124
DF	2
Chi Square	3.985
Chi Square P-Value	.1364
G-Squared	•
G-Squared P-Value	•
Contingency Coef.	.160
Cramer's V	.162

Observed Frequencies for FINAL CK5 (AT 1%), GRADE
Row exclusion: DCIS STATVIEW DATASET

	HG	IG	LG	Totals
0	73	46	10	129
1	17	5	0	22
Totals	90	51	10	151

Summary Table for FINAL CK5/6 (AT 1%), GRADE
Row exclusion: DCIS STATVIEW DATASET

Num. Missing	125
DF	2
Chi Square	4.519
Chi Square P-Value	.1044
G-Squared	4.680
G-Squared P-Value	.0963
Contingency Coef.	.171
Cramer's V	.174

Observed Frequencies for FINAL CK5/6 (AT 1%), GRADE
Row exclusion: DCIS STATVIEW DATASET

	HG	IG	LG	Totals
0	56	42	5	103
1	34	11	2	47
Totals	90	53	7	150

Summary Table for FINAL CK14 (AT 1%), GRADE
Row exclusion: DCIS STATVIEW DATASET

Num. Missing	119
DF	2
Chi Square	4.325
Chi Square P-Value	.1150
G-Squared	4.658
G-Squared P-Value	.0974
Contingency Coef.	.164
Cramer's V	.167

Observed Frequencies for FINAL CK14 (AT 1%), GRADE
Row exclusion: DCIS STATVIEW DATASET

	HG	IG	LG	Totals
0	69	49	7	125
1	23	6	2	31
Totals	92	55	9	156

Summary Table for FINAL KI67 (AT MEDIAN, 5%), GRADE
Row exclusion: DCIS STATVIEW DATASET

Num. Missing	121
DF	2
Chi Square	13.981
Chi Square P-Value	.0009
G-Squared	14.052
G-Squared P-Value	.0009
Contingency Coef.	.288
Cramer's V	.301

Observed Frequencies for FINAL KI67 (AT MEDIAN, 5%), GRADE
Row exclusion: DCIS STATVIEW DATASET

	HG	IG	LG	Totals
0	26	28	8	62
1	65	24	3	92
Totals	91	52	11	154

Summary Table for FINAL MCM2 (AT MEDIAN, 27.5%), GRADE
Row exclusion: DCIS STATVIEW DATASET

Num. Missing	117
DF	2
Chi Square	21.592
Chi Square P-Value	<.0001
G-Squared	21.981
G-Squared P-Value	<.0001
Contingency Coef.	.347
Cramer's V	.370

Observed Frequencies for FINAL MCM2 (AT MEDIAN, 27.5%), GRADE
Row exclusion: DCIS STATVIEW DATASET

	HG	IG	LG	Totals
0	28	36	7	71
1	66	17	4	87
Totals	94	53	11	158

Summary Table for ARCH, GRADE

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	106
DF	6
Chi Square	47.715
Chi Square P-Value	<.0001
G-Squared	•
G-Squared P-Value	•
Contingency Coef.	.469
Cramer's V	.376

Observed Frequencies for ARCH, GRADE

Row exclusion: DCIS STATVIEW DATASET

	HG	IG	LG	Totals
CRIB	25	33	10	68
MICROPAP	11	4	3	18
PAP	3	10	0	13
SOLID	59	11	0	70
Totals	98	58	13	169

Summary Table for NECROSIS, GRADE

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	106
DF	6
Chi Square	69.559
Chi Square P-Value	<.0001
G-Squared	•
G-Squared P-Value	•
Contingency Coef.	.540
Cramer's V	.454

Observed Frequencies for NECROSIS, GRADE

Row exclusion: DCIS STATVIEW DATASET

	HG	IG	LG	Totals
MARKED	39	3	0	42
MILD	22	24	1	47
MOD	28	10	0	38
NONE	9	21	12	42
Totals	98	58	13	169

Summary Table for CI, GRADE

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	108
DF	6
Chi Square	50.846
Chi Square P-Value	<.0001
G-Squared	•
G-Squared P-Value	•
Contingency Coef.	.483
Cramer's V	.390

Observed Frequencies for CI, GRADE

Row exclusion: DCIS STATVIEW DATASET

	HG	IG	LG	Totals
MARKED	37	0	0	37
MILD	27	28	4	59
MOD	24	11	1	36
NONE	10	17	8	35
Totals	98	56	13	167

Table 22: Paired Univariate Analysis for Histological Grade.

3.9.10 Assignment of Molecular Subtypes to DCIS

3.9.10.1 Molecular Subgrouping of DCIS

Surrogate molecular subgroups (76) were assigned to the pure DCIS cohort using the criteria in Table 23. These groups broadly correlate to the five molecular subtypes proposed by Perou et al. A subgroup which was both HER2 and ER positive was also identified within the series and categorised as luminal/HER2.

DCIS subgroup	IHC Expression Pattern	Number of Cases	% of Cases
Luminal A	ER + PR+ HER2 -	103	38
Luminal B	ER + PR- HER2 -	55	20
HER2	ER – HER2 +	47	17
Basal-like	EGFR+ and/or CK5, CK5/6, CK14+	50	18
Luminal HER2	ER + HER2 +	20	7

Table 23: Definitions for Surrogate Molecular Subgroups for DCIS:

Additionally, the expression of Ki67 was examined in the ER positive and HER2 negative DCIS cases in order to define luminal A and luminal B according to proliferation rate (Table 24). Subgroups of luminal A1 and luminal B1 were defined as ER positive, Ki67 low and ER positive, Ki67 high respectively, in addition to the alternative method of defining cases according to ER and PR status, i.e. ER positive,

PR positive as luminal A and ER positive, PR negative as luminal B. Table 24 gives details of numbers of ER positive cases as assigned by high or low Ki67 expression.

Group	Definition	No. Cases (TMA total n = 280, relevant (ER+) cases n= 158)	% of assessable Cases
Luminal A1	ER+ Ki67 low (<5% tumour cells)	78	49
Luminal B1	ER + Ki67 high (>5% tumour cells)	80	51

Table 24: Subgrouping of ER positive DCIS with Ki67 expression; cut-off defined at median, 5%.

3.10 Immunophenotypic Subtypes of DCIS and Recurrence

Three HER2 positive (HER2 group) cases recurred (n=47; 6.4%) and three of the basal-like group (n=50; 6%) had a recurrence (Table 20). In this cohort of pure DCIS the luminal A group (n=103), as defined above, showed no association with recurrence; nine of the recurrent cases belonged to the luminal A group. Only one of the recurrent cases was found in the luminal B group (n=55).

For ER positive cases (n=95), 51% (n=48) had high Ki67 expression. High Ki67 status resulted in a 19% (n=18) shift from the luminal A (ER positive but low proliferation, defined as A1) to the luminal B1 (ER positive but high proliferation, defined as B1) group. In this series, seven of the luminal B1 group and four of the luminal A1 had a

recurrence. Overall, cases of DCIS that were ER and PR positive had a higher incidence of recurrence than cases which were ER positive with high Ki67 expression. In our series 74% of DCIS are ER positive. Approximately 70-80% of invasive ductal carcinomas of no special type (NOS) are ER positive according to the literature (188). This is related to cytonuclear grade of DCIS in the present series (p value < 0.001); 61% of high grade cases (HG) are ER positive (n=90), 92% of intermediate grade (IG) (n=50) and 100% of low grade (LG) DCIS are ER positive n=9). Previous studies have demonstrated variation in ER expression in high grade DCIS, ranging from 31.8% (189), 69% (190) to 72% (191).

3.11 Discussion

The genomic profiling of invasive breast cancers has revealed subtypes of invasive breast cancer, namely luminal A, luminal B, HER2, basal, normal and (subsequently) null. A search of the literature (Pubmed in 2014: search criteria; DCIS + molecular subtype, DCIS + IHC, DCIS + basal, DCIS + luminal, DCIS + phenotype) has yielded a small number of studies that examine the possibility of using IHC as a surrogate for genetic profiling to classify DCIS in a similar manner. To date only a few (82, 192) of these have taken a comparative approach by assigning the genetic molecular profile to samples then assessing the corresponding IHC as opposed to simply assigning groups by IHC expression alone.

In this series immunohistochemistry (IHC) on “pure” DCIS was performed to ascertain if any potential marker or markers could be identified as a biological indicator of known progression to either a recurrence of DCIS or invasive disease. However, no association or predictor has been found for recurrence of DCIS with any of ER, PR, HER2, EGFR or the basal cytokeratins 5, 5/6 or 14, high or low expression of proliferative markers in univariate analysis. Unsurprisingly, a correlation is seen

between cytonuclear grade increased expression of proliferative markers (Ki67 and MCM2), which largely mirrors findings in invasive breast carcinoma.

None of the IHC biomarkers evaluated in this cohort were significantly associated with recurrence of disease, although this may be due to the small numbers of patients in this series who suffered with recurrence. Additionally, the cohort for this study was based upon patients diagnosed with DCIS between 1995 and 2006 with 50% (n=71) of patients undergoing a mastectomy within 65 days of original histology report. Thomas et al (193) examined data from the Sloane Project, currently the largest prospective audit on DCIS. Examination of over 8000 cases indicates that mastectomy rates dropped to 26% between April 2003 and March 2012 with more women undergoing breast conserving surgery (BCS). They also highlight the significant degree of practice variation between institutions affecting patient outcome. These differences in practice should be considered when looking at data sets from an individual institution.

Molecular profiling of *invasive* breast carcinomas gives subgroup frequencies of luminal A - 40%, luminal B - 20%, HER2 - 10-15% and basal-like - 15-20%. (76, 82, 194, 195). In this series of DCIS the subgroup frequencies for DCIS as determined by IHC profiling were luminal A - 37%, luminal B - 20%, HER2 - 17%, luminal/HER2 - 7% and basal-like - 18%, indicating an essentially similar frequency between the invasive carcinoma immunosub-types and that seen in DCIS.

There is, however, no consensus agreement on assigning sub-type based on IHC panels for invasive disease, and this has not been satisfactorily addressed for DCIS. Whilst there is a increase in the frequency of HER2 positive DCIS compared to invasive breast carcinoma, 7% of these would be assigned to the luminal A group based on the surrogate panels suggested by some authors (82, 196, 197). In this cohort these patients are designated to a small group that is positive for both HER2 and ER as thus as luminal/HER2.

Other studies have looked at the possibility of using IHC as a method of classifying DCIS, either to define potential molecular subtypes (86, 196-199) or specifically to identify a sub-set of basal-like DCIS, or to predict outcome or recurrence. Tamimi et al (200) identified subgroups in both invasive breast carcinomas (n=2897) and DCIS (n=272) noting that the luminal A group was more prevalent in the invasive cohort, whilst the HER2 and luminal B groups had a higher incidence in DCIS. Zhou et al (198) identified a basal-like subtype of DCIS with an increased risk of recurrence of disease. They constructed TMAs from 485 cases of primary DCIS found in Swedish women between 1986 and 2004. Using IHC markers (ER, PR, HER2, CK5/6 and EGFR) as surrogates for genomic profiling they defined groups as triple negative, non-triple negative and basal-like. Results for 392 women were available. Basal-like DCIS in their series (n=32) had a higher risk of local recurrence with hazard ratio (HR) of 1.8 (Confidence interval (CI) 95% 0.8 – 4.2) and also invasive risk recurrence of 1.9 (CI 0.7 – 5.1) but these results were not statistically significant. They found no significantly increased risk of recurrence of disease for patients with triple negative DCIS.

Meijnen et al (196) evaluated 16 IHC markers identifying six (ER, PR, androgen receptor, Bcl2, HER2 and p53) which they suggested were suitable for defining DCIS into two major groups, luminal and non-luminal, and 5 subgroups. They also noted that intermediate grade DCIS had more IHC similarities to well differentiated than poorly differentiated disease. Clark et al (197) also classified DCIS into molecular subgroups with IHC, suggesting a role for BCL2 as a possible identifier for good prognosis. Overall, therefore the present, and other, studies show that it is possible to identify differences between cases of DCIS using panels of IHC markers. Such markers show correlations with each other and with features of known clinico-pathological relevance and can be combined into sub-groups equivalent to those in invasive breast cancer. However, no individual marker, or combination, shows a

strong association with recurrence of disease and thus to be of significant clinical value at this time.

Few studies have looked directly at genomic and IHC parameters in the same lesions, either invasive or in situ. Nielsen et al (82) studied *invasive* breast carcinoma with both IHC and genomic profiling and identified a concordant set of basal-like tumours using four IHC markers (HER2, EGFR, CK5/6 and ER). In the present series we have defined a basal-like group by utilising three basal cytokeratin antibodies (CK14, CK5 and CK5/6) in addition to EGFR, ER, PR and HER2, but we note that the frequencies of expression of CK5/6 and CK5 showed distinct disparity. This highlights one of the concerns in the definition of molecular subtypes using IHC; the variability of clones and the interpretation of scoring are parameters that currently have no defined values. What constitutes a positive CK score varies in the literature, with some researchers advocating any staining as positive (196, 200) and some using a threshold of $\geq 10\%$ (197, 199). Indeed, one of the issues that requires addressing before a molecular taxonomy equivalent to genomic profiling, for either invasive or DCIS, can be adopted by application of IHC is the precise and globally acceptable definitions using specific antibodies, and possibly even clones. Robust scoring criteria and a need for standardisation of collection, fixation, processing and reporting are also fundamental to accurate assignation of molecular subtypes. Tang et al (201) reviewed the IHC classification of invasive breast cancers and advocated a molecular classification, which could be similarly used for DCIS. They suggest for invasive breast cancer subgrouping as luminal A (ER positive, HER2 negative), luminal B (ER positive, HER2 positive), HER2 group (ER negative, HER2 positive), basal subgroup (ER negative, HER2 negative, CK5/6 and/ EGFR positive) and an unclassified group (ER negative, HER2 negative, CK5/6 negative, EGFR negative).

The present study aimed to identify the possibility of using IHC as a genomic surrogate to identify subgroups of DCIS and to investigate associations with patient outcome. We have found that molecular subgroups can be identified using an IHC approach. However, further work is required to evaluate the usefulness of such a system in DCIS, particularly with regard to prediction of recurrence. Retrospective studies will probably lack the statistical power or accuracy required to give reliable data, given the variation in clinical management in historical series, including the proportion of cases treatment by breast conserving surgery, differences in the definition of a tumour-free margin of excision, as well as the use of radiotherapy and hormone treatments.

Chapter 4. DCIS Associated with Invasive Breast Carcinoma

4.1 Introduction

The molecular profiling and immunohistochemical studies described in Chapter 3 focused upon the IHC marker expression and the propensity for “pure” DCIS to become invasive or recur as further DCIS. However, DCIS may also be expressed alongside an invasive breast lesion. Previous genomic studies have demonstrated that similar genetic profiles exist between DCIS and synchronously expressed invasive breast disease (202-204). Conversely, there is some evidence of intra-tumour heterogeneity between some DCIS and the respective invasive breast lesion given rise to the possibility of divergent clonal pathways within an individual tumour (202). This may indicate that, in synchronous DCIS and invasive breast disease at least, that DCIS is already at an advanced stage committed to invasive progression. We aimed therefore to explore the possibility of different IHC profiles or frequencies in a subset of DCIS with associated invasive carcinoma and to determine whether any potential difference existed between pure DCIS and DCIS lesions associated with invasive breast tumours. These would be stained with the same antibody panel as the pure DCIS in Chapter 3. This subset would also provide part of the cohort for further genomic studies (see Chapter 5).

4.2 Materials and Methods

DCIS and associated invasive tissue were provided by King’s Health Partners Tissue Bank (KHPB) with tissue microarrays constructed in the same manner as described in section 2.2. Similarly, TMA sections were stained with the same antibodies using

the similar conditions as set out in section 3.2. However, in this instance cores were taken from a minimum of two defined processes, namely DCIS and invasive components. For DCIS regions three 2mm cores were taken and for the invasive component three 0.6mm cores were taken. The DCIS component was identified by the presence of a complete basement membrane and myoepithelial layer surrounding the malignant epithelial cell islands by the author and consultant breast pathologist (SEP). In lesions where there was a doubt regarding the presence of an intact membrane, IHC for the marker smooth muscle myosin heavy chain (SMMHC) was used to identify DCIS by demonstration of the surrounding myoepithelial cell layer (Figure 30).

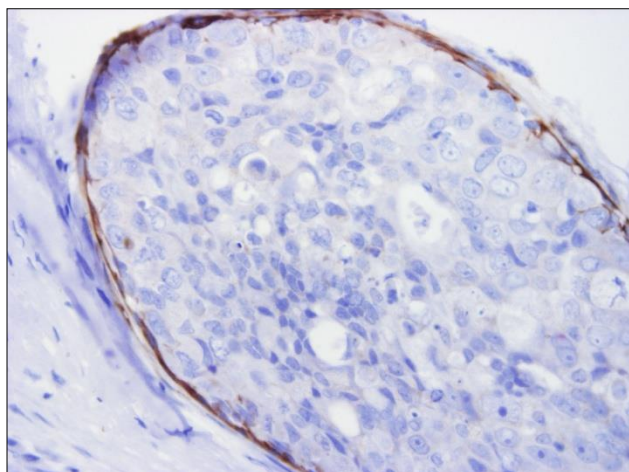


Figure 30: SMMHC IHC showing intact myoepithelial cell layer around DCIS.

A total of 41 cases dating between 1989 and 2004 with synchronous invasive breast disease and DCIS lesions was identified and stained with ER, PR, HER2, EGFR, CK5/6, CK14, MCM2 and Ki67. IHC scoring was carried out as per section 3.7.

Molecular subgroups were assigned as defined by Perou et al using the following criteria: ER positive, PR positive, HER2 negative = luminal A; ER positive, PR negative, HER2 negative = luminal B; HER2 positive, ER negative = HER2 group; ER

negative, PR negative, HER2 negative with any/or EGFR, CK5/6, CK14 = basal-like; no expression of any of ER, PR, HER2, basal CKs or EGFR = null. A subgroup of disease which was HER2 positive and ER positive was also noted within our series and designated luminal HER2. Additionally, the expression of proliferative markers was examined in ER positive cases to determine any association with prognosis and recurrence. Subgroups luminal A1 and luminal B1 were defined as ER positive, Ki67 low and ER positive, Ki67 high respectively as opposed to ER positive, PR positive (luminal A) and ER positive PR negative (luminal B) within the HER2 negative subset.

4.3 Results

4.3.1 Individual IHC Antibody Staining Concordance between Invasive Tumour and DCIS in the Same Case:

The TMAs constructed and stained with IHC markers had scores generated for both the DCIS and invasive carcinoma components within each case. The DCIS and invasive component scores with each antibody were compared. Table 25 shows the association between DCIS and invasive IHC scores.

Antibody IHC scoring method and cut-off value)	Positive in DCIS and invasive carcinoma	Positive in invasive only	Positive in DCIS only	Negative for both	DCIS loss of core	Invasive loss of core	Total
ER (Allred >3)	23	3	2	2	11	0	41
PR (Allred >3)	9	1	3	8	20	0	41
HER2 (HER2 score 3+)	8	0	7	13	13	0	41
EGFR (HER2- like score 3+)	5	6	2	4	24	0	41
CK5/6 (>2% tumour cells stained)	1	4	0	32	4	0	41
CK14 (>2% tumour cells stained)	1	4	2	30	4	0	41
Ki67 (>5% tumour cells stained)	9	16	1	10	3	2	41
MCM2 (>5% tumour cells stained)	9	1	6	2	20	3	41

Table 25: Comparison of IHC scores between cases of admixed DCIS and invasive carcinoma

4.3.2 Assignment of Molecular Subgroups

Molecular subgroups were assigned according to both the invasive carcinoma IHC scores and DCIS IHC scores. Tables 26 and 27 give details of numbers and subgroups and concordance. Table 26 shows the number of cases assigned to each group according to DCIS or Invasive component. Table 27 shows the distribution according to proliferative (Ki67) and hormonal status.

	Invasive Carcinoma	DCIS	Paired Cases (Invasive carcinoma and DCIS match in same case)	Unpaired cases (Invasive carcinoma and DCIS do not match in the same case)
Luminal A (ER positive, PR positive, HER2 negative)	14	6	4	2
Luminal B (ER positive, PR negative, HER2 negative)	3	4	3	0
HER2 (HER2 positive, ER negative)	0	4	1	3
Basal-like (ER negative, PR negative, HER2 negative, EGFR /CK5/ CK5/6 and/or CK14)	12	9	4	5
Luminal/ HER2 (ER positive, HER2 positive)	7	11	7	0
Null (All negative)	1	1	0	1
Total	37	35	18	12

Table 26: IHC Molecular Subgroups for Invasive Carcinomas and Associated DCIS

	Invasive Carcinoma	DCIS	Paired Cases (Invasive carcinoma and DCIS match in same case)	Unpaired cases (Invasive carcinoma and DCIS do Not match in the same case)
Luminal A1 ER positive, Ki67 high expression	10	9	6	3
Luminal B1 ER positive, Ki67 low expression	10	16	6	4
Total	*20	25	12	7
	*3 cases no Ki67 Score			

Table 27: IHC Molecular Subgroups for Invasive Carcinomas and Associated DCIS

for ER Positive and Ki67 Expression

4.4 Discussion

In this series of DCIS associated with invasive breast carcinoma several observations can be made. For ER positive DCIS and invasive breast disease there is a high degree of concordance with 83% (n=25/30) of cases showing a similar ER status. Analysis of the Allred scores also shows that cores have a similar degree of ER expression. A similar pattern is seen with PR, with 81% (n=17/21) having concordant scores.

HER2 expression with IHC showed 75% concordance (n=21/28); although 25% of the DCIS was HER2 IHC positive whereas the invasive breast disease was negative, even in this small series.

For EGFR there was no distinct pattern of correlation between DCIS and invasive breast disease; 3 cases were EGFR positive in both the invasive breast cancer and the DCIS, 3 cases showed positivity in the invasive disease only and 2 cases positivity in the DCIS only.

Interestingly, there was also little similarity between invasive breast disease and DCIS with regard to basal cytokeratin expression; in 5 cases positive CK5/6 expression was seen in the invasive carcinoma whilst the DCIS was CK5/6 negative. Similarly, CK14 positive invasive breast disease (n=4) was negative in the associated DCIS in 3 cases. Conversely, DCIS was positive for CK14 in 3 cases and negative in the associated invasive cancer in two. One DCIS case was CK5/6 positive but negative in the invasive breast disease.

Ki67 showed concordance (scores within 5% of each other) for 63% of admixed DCIS and invasive breast disease (n=22). Of note, the 37% (n=13) of discordant cases had elevated Ki67 in the invasive breast disease in all cases (range = 7-32% mean = 10%). MCM2 expression was concordant in the DCIS and the invasive disease in 37% of cases (n=7/19), 52% (n=10/19) of cases showed elevated MCM2 expression

(range = 7-73% mean = 38%) in the invasive tumour compared to the DCIS and 11% (n= 2/19) showing elevated MCM2 in the DCIS lesion (range= 12-49% mean =30%). Unlike the pure DCIS cohort, HER2 expression was associated with ER positivity in the DCIS component with 73% (n=11/15) showing positive staining for both. In the invasive component, unexpectedly, 100% of the HER2 positive disease was ER positive (n=7/7). Although there are many studies on ER positive HER2 negative breast cancer there are few articles addressing ER positive HER2 positive lesions. Ryden et al (205) describe HER2+ and ER+ breast cancer as a separate subgroup of tumours with poor prognosis in premenopausal breast cancer. Alqaisi et al (206) found co-expression of ER-positive HER2 positive disease was found in younger age, had higher grade and increased visceral involvement in de novo metastasis.

Unfortunately, in the DCIS TMA there was a marked loss of cores for ER, PR, EGFR, HER2 and MCM2 during the antigen retrieval step. The reason for this remains unclear as all TMAs were constructed with the same guidelines as other TMAs, including the invasive TMA from the same tissue samples, which retained tissue during antigen retrieval steps. The same cases for each antibody had tissue loss possibly indicating a fixation or necrotic tissue problem. It was possible therefore to assign only 35 cases to molecular subgroup for both the invasive and the in situ elements..

Intriguingly, however, in this cohort such molecular subgrouping does not show perfect concordance between the DCIS and the associated invasive tumour for most groups. Only the luminal/HER2 group has complete concordance with all available samples being ER positive and HER2 positive in both the DCIS and HER2. 66% of luminal A, 0% of luminal B, 44% of basal-like and 0% of null had concurrent IHC staining in this series. This could be due to several factors:

1. Tumour heterogeneity and, in particular, clonal expansion of particular cell types once invasion has occurred;

2. Synchronous development of unrelated tumour and DCIS within the same patient (highly unlikely);
3. Down- or up-regulation of proteins once invasion has begun.

The limitations of this experiment must be fully acknowledged before such conclusions can be drawn. The substantial loss of cores in the DCIS TMA during IHC staining has compromised the ability to group samples by more than one marker and have reduced the cohort size significantly.

Chapter 5. Genomic Studies on DCIS

5.1 Introduction

Understanding the genomic basis of DCIS and how it relates to invasive breast carcinoma may help elucidate the underlying mechanisms of progression to invasive disease or demonstrate that DCIS in certain instances may already represent a commitment to an invasive phenotype. The progression of DCIS to an invasive phenotype has been previously studied using gene expression profile analysis (203, 207) and array comparative genomic hybridisation (aCGH) in small numbers of studies and cases (208-210). These studies have demonstrated that DCIS lesions of the same grade have similar genetic profiles to the concurrent invasive tumour, although qualitative differences such as prevalence are reported in matched series (211).

Copy number variations (CNVs) are genomic variations of greater than one kilobase affecting the number of copies of genes possessed by an individual compared to a reference sample. First reported by Charles Lee in 2002 (212) CNVs represent deletions, insertions or duplications of DNA. Shlien et al (213) described the importance of CNVs succinctly:

“DNA copy number variations (CNVs) are an important component of genetic variation, affecting a greater fraction of the genome than single nucleotide polymorphisms (SNPs). The advent of high-resolution SNP arrays has made it possible to identify CNVs. Characterization of widespread constitutional (germline) CNVs has provided insight into their role in susceptibility to a wide spectrum of disease, and somatic CNVs can be used to identify regions of the genome involved

in disease phenotypes. The role of CNVs as risk factors for cancer is currently underappreciated.”

Identification of CNVs in DCIS may provide clues to progression and help improve targeted treatment for both those with low and high risk invasive potential.

5.2 Genetic Aberrations of DCIS and Invasive Breast Disease – Literature Review

Changes in DNA copy numbers may be inherited or due to somatic mutations. Copy number variants (CNVs) usually refer to inherited variation whilst copy number aberrations (CNAs) refer to somatic mutation (214). Whilst beyond the scope of the thesis to identify every potential copy number gain or loss found in invasive breast disease in the literature, due to the vast array of data, commonly found copy number aberrations (CNAs) have been identified (215-219). These include gains on 1q, 6p, 8q, 11q, 16p, 17q, 19, 20q and losses on 5q, 6q, 7p, 13q, 16q, 17p and 22q. In DCIS, several studies have identified a variety of CNAs (202, 207, 210, 220, 221). Copy number gains on 1q, 5q, 8q, 16p, 19q, 20 and losses on 1p, 3q, 6q, 8p, 9p, 11q, 13q, 16q, 17p and 17q are reported. These include known loci for *HER2* (17q12), ER (6q25.1), EGFR (7p11) and p53 (17p1).

In addition to the known CNAs there are many novel potential candidate genes identified in both DCIS and invasive breast disease that could be biomarkers for therapeutic intervention. Hanneman et al (192) identified 35 genes which differed between DCIS and invasive breast disease and 43 genes that differed between low grade and high grade DCIS in 2006. Since then a few studies (210, 222-224) have identified genes that may be implicated in transformation of DCIS into an invasive phenotype.

5.3 Materials and Methods

Several criteria were used for selections of MIP array samples. Due to the cost, it was deemed unfeasible to perform analysis on the whole cohort from the DCIS TMA set. Additionally, although only a comparatively small amount of tissue was required from the DCIS samples there was an issue of tissue availability. In general DCIS is not markedly cellular, TMAs had already been manufactured from the cases, and clearly one has a commitment not to exhaust a resource that must legally be retained as part of the patients' diagnostic record. Subsets for genomic analyses were therefore chosen to represent two groups, namely pure DCIS and DCIS associated with Invasive component. To further represent the larger cohort, DCIS lesions were chosen with the following IHC profiles: ER positive, HER2 positive and triple negative. Where a DCIS lesion was matched to invasive tissue samples were taken from both components and if possible a sample of adjacent normal tissue. Table 28 gives the number and type of sample selected for MIP array analysis.

Specimen type	Number of Cases
Triple negative pure DCIS	12
Triple negative DCIS associated with invasive breast disease* and	10
Triple negative invasive breast disease associated with DCIS*	10
ER positive pure DCIS	10
ER positive DCIS associated with invasive breast disease^ and	9
ER positive invasive breast disease associated with DCIS^	9
HER2 positive pure DCIS	8
Matched normal from ER positive DCIS and invasive breast disease^	7
Matched normal from triple negative DCIS and invasive breast disease*	8
Total number of samples	83

Table 28: Sample Selection and Case Numbers for MIP Array Analysis.

[* and ^ represent different microdissected components of the same cases]

5.4 Hypothesis for MIP Arrays Studies in DCIS and Invasive Breast Carcinoma

Genomic analysis of tissue samples may provide additional knowledge in our understanding of breast cancer progression and help define treatments in the future. Whilst not exhaustive in sample size this does represent one of the largest studies on DCIS using MIP array to date. A comparison of the various groups may shed light on genomic differences found within the different disease states i.e. precursor versus invasive breast cancer. This could be used to determine if the expression the immunohistochemical markers used in Chapter 3 are present at the genomic level (as expected), for example, is there HER2 amplification in the HER2 positive IHC cases? It also creates a database of genomic data with to study alongside the immunohistochemical and demographic data. This database will provide a resource for further research beyond the scope of this thesis. The analysis will also provide insights into variations between genomic changes seen in DCIS and invasive breast cancers and may present novel genes not currently identified in the literature. Identification of genes that represent an underlying commitment to invasive potential established before the expression of a molecular subtype may also be found. To address these issues comparisons of different comparative analysis of the following groups was carried out to help elucidate possible genetic aberrations, for all cases of DCIS, both pure and that associated with invasive disease:

1. ER+ DCIS compared to non ER+ DCIS
2. ER+ pure DCIS compared to ER+ DCIS associated with invasive disease
3. ER+ pure DCIS compared to ER+ invasive breast disease
4. All ER+ DCIS (pure and matched) compared to all ER+ invasive disease
5. Triple negative pure DCIS compared to triple negative invasive breast disease
6. All triple negative DCIS compared to all non-triple negative DCIS

7. Triple negative pure DCIS compared to triple negative DCIS associated with invasive disease
8. All triple negative DCIS compared to triple negative invasive disease
9. HER2+ pure DCIS compared to non-HER2+ DCIS

5.4.1 DNA Extraction

DNA was extracted using a modified technique incorporating elements from the QIAGEN FFPE DNA kit (QIAGEN Ltd. Manchester UK) and protocols provided by Dr Jorge Reis-Filho (Cheaty Beatty Laboratories, London, UK). 4µm sections were cut from blocks already identified from the TMA cohorts, and H&E stained to ascertain the amount of available tissue. Where a suitable quantity of DCIS and/or invasive tumour/normal tissues was present, up to 20 sections were cut at 8µm and dried overnight at 37°C. Sections were dewaxed in xylene then stained with nuclear fast red dye. Sections were viewed through a dissecting microscope and any areas of tissue not required for DNA analysis were scraped away using a fine aspirate needle (gauge 10). This included any stroma, lipid or connective tissue. Sections were immersed in sodium thiocyanate overnight to remove formalin fixation crosslinks. Slides were then rinsed in changes of ultra-pure H₂O and sections examined under the dissecting microscope where components such as DCIS, invasive tumour and normal tissues were scraped into individual 1.5ml eppendorf tubes using individual aspirate needles for each component. Samples were immersed in proteinase K for 48 hours at 40°C to aid proteolysis, followed by DNA extraction using the QIAamp DSP DNA FFPE Tissue Kit (Qiagen, Manchester UK). DNA quality was initially assessed by taking a spectrophotometer reading (Nanodrop, Thermo Scientific Wilmington, Delaware, USA) to determine DNA presence. DNA absorbs light at 260nm, a spectrophotometer can detect the amount of light absorbed in a sample, thus the more DNA the greater the absorbance of light. However, contaminants such

as proteins and RNA may also absorb light at 260nm so errors are possible. Once the presence of DNA was confirmed by this approach, samples were then assessed using a fluorimeter (Qubit- Life Sciences, Paisley, UK). The Qubit system uses fluorescent dyes that bind to target molecules (e.g. DNA); these dyes have a minimal fluorescent signature which becomes greatly enhanced upon target binding. When compared to a standard sample of known concentration an accurate DNA concentration can be given. Protocols for DNA extraction are listed in Appendix 6.

5.4.2 MIP Array

DNA samples were shipped to Affymetrix MIP laboratory (California, USA). Samples were allocated an anonymised number and all identifying data removed to ensure sample blinding, including the matched normal tissues. Using the OncoScan™ FFPE Express Service samples underwent MIP array assays.

Data quality was assessed using the sample 2-point relative standard error. The majority (95%) of FFPE tumours samples applied to the MIP arrays passed the 2p-RSE threshold (GSE31424).

5.4.3 Statistical methods

Tumour Aberration Prediction Suite (TAPS), a bioinformatics tool for the identification of allele-specific copy numbers in tumour samples using data from Affymetrix SNP arrays, was run on all samples to obtain allele specific absolute copy numbers at genomic segment level. Raw output tables from TAPS output are seen on the attached disc. Gene centric versions of these tables i.e. each row represents a gene and its respective CN and other parameters are directly copied from TAPS raw tables are also provided (attached disc). This shows gene level copy numbers for each sample. Using these gene centric tables, intra-group gene centric count/percentage tables were created according to the IHC groups (ER+, triple negative and HER2+) (also

see attached disc). These tables show counts/percentage for each gene and each CN type, as well as the respective sample IDs.

To compare groups, Fisher's exact tests were performed between two groups for each copy number type. For example, frequencies were calculated for "gains" for each segment among the two groups into a contingency table which was used to run Fisher's test, resulting in a p-value for each segment for that copy number type. Files are stored on the attached disc and give Fisher tests output for all segments for each comparison respectively. Genome frequency plots for these tables are also provided (see attached disc). Significantly different segments between 2 groups i.e. based on p-values cut-off of ≤ 0.05 , are also presented on the attached disc, along with tables and genomic frequency plots for segments based on the above mentioned threshold.

5.4.4 Analyses of Raw Data Tables

The raw data (TAPS tables) produced a significant amount of data, which is too large for a single thesis to adequately analyse. Mining of the data to elicit possible key genes was determined to be a suitable way forward; this approach has been used by others previously (221).

Data was compiled into groups for comparison e.g. ER positive pure DCIS compared to pure non ER positive DCIS (similarly all triple negative, HER2 etc.).

Genomic regions were sorted into those with significant p-values for the various genomic states: amplification, deletion, copy neutral loss of heterozygosity (cnLOH), copy deletion loss of heterozygosity (cdLOH) and total loss.

Of the samples with significant p values, genomic bands with the greatest difference between the number of samples (and the greatest percentage difference) were selected, e.g. in ER positive pure DCIS there is genomic band (region) where 5 samples had amplification within this region compared to no amplification in any of the pure non ER positive samples.

In many instances there were multiple genetic aberrations listed in a general locus of a particular chromosome or arm. In others there were only one or two genetic aberrations identified. A rationale that the occurrence of the same genetic aberrations found in multiple samples are more likely to be key players than genetic aberrations found in any one individual sample representing heterogeneity or genetic plasticity in tumourigenesis was adopted. Although this concept may be potentially flawed, analysis of multiple genes on the same locus also produces very complex and varied biological processes when analysed by PANTHER (see 5.2.2).

In addition to the above method, a literature search of genes associated with DCIS and invasive progression previously identified as potential drivers by other researchers was carried out.

5.5.5 Analysis of MIP array Data Using “PANTHER” Classification System

The large amount of genomic data generated, and subsequently analysed by the Breakthrough Breast Cancer/Research Oncology, King’s College London bioinformatics team, based upon the MIP array results requires further analysis in order to identify similarities and differences between groups. There are many gene analysis software packages available online (e.g. GeneOntology, Webgestalt, GeneGo, Pubgene, TopGeneSuite). It is beyond the scope of this thesis to determine the pros and cons of each individual package. Under advice from the Breakthrough Breast Cancer/Research Oncology, King’s College London bioinformatics team, and based upon the type of analysis required, the “protein annotation through evolutionary relationships” or PANTHER protein and gene analysis software was selected. In 2013 Mi et al (225) described the analysis software thus: “PANTHER (protein annotation through evolutionary relationship) classification system (<http://www.pantherdb.org/>) is a comprehensive system that combines gene function, ontology, pathways and

statistical analysis tools that enable biologists to analyse large-scale, genome-wide data from sequencing, proteomics or gene expression experiments.”

5.5.6 Analysis of DCIS and Invasive Breast Disease Subgroups.

Groups were compared using the 7 copy number types (excluding undefined segments) shown below. Generalized copy number types of "gains" and "losses" were also included (Table 29).

Cn	mCn	New CN Type	Combined CN Type
NA	NA	Undefined	Undefined
0	0	Total Loss	Total loss
2	NA	Normal with unknown allele specificity	Normal
2	1	Allele specific normal	
1	NA/0	Copy deletion LoH	Copy deletion LoH
2	0	Copy neutral LoH	Copy neutral LoH
3	NA	Single copy gain with unknown allele specificity	Single copy gain
3	0	Copy neutral single copy gain	

Table 29: Copy number analysis types for MIP arrays.

5.8 Results

5.8.1 DNA Extraction for MIP Arrays

For each of the 82 samples undergoing MIP array DNA quantities and Qubit fluorescence readings can be found in appendix 7.

5.8.2 MIP Array Maps

All MIP array data was analysed by the Bioinformatics Unit of King's College London, Breakthrough Breast Cancer. MIP array data was analysed by three Independent researchers (AG, HM & SR). A total of 73 cases (Table 30) were provided for analysis separated into the different phenotypic groups. These phenotypic groups were then compared with each other (Table 31).

All MIP array chromosomal maps can be found in appendix 8 & 9. Representative figures are given below for ER positive pure DCIS (Figure 31), HER2 positive pure DCIS (Figure 32) and triple negative pure DCIS (Figure 33). Additionally, representative figures showing comparisons of normal breast tissue, DCIS and invasive breast disease for ER positive (Figure 34) and triple negative cases (Figure 34) are given below.

Specimen type	Number of Cases
ER positive pure DCIS	8/10
ER positive DCIS associated with invasive disease	9/9
ER positive invasive breast disease associated with DCIS	9/9
Triple negative pure DCIS	9/12
Triple negative DCIS associated with invasive breast disease	10/10
Triple negative invasive disease associated with DCIS	7/10
HER2 positive pure DCIS	7/8

Table 30: MIP array cases with chromosomal maps

Title	Group1	Group2
ER+ DCIS compared to non ER+ DCIS	ER Pure DCIS, ER DCIS with invasive disease (n=17)	TN pure DCIS, TN DCIS with invasive disease, HER2 positive pure DCIS (n=26)
ER+ DCIS associated with invasive breast disease compared to ER+ pure DCIS	ER+ DCIS with invasive disease (n=9)	ER+ pure DCIS (n=8)
ER+ invasive breast disease compared to all ER+ DCIS	ER+ invasive breast disease (n=9)	ER+ pure DCIS, ER DCIS with invasive disease (n=17)
ER+ pure DCIS compared to ER+ invasive breast disease	ER+ pure DCIS (n=8)	ER+ invasive breast disease (n=9)
Triple negative invasive breast disease compared to all triple negative DCIS	TN invasive breast disease (n=7)	TN Pure DCIS, TN DCIS with invasive disease (n=19)
Triple negative DCIS associated with invasive breast disease compared to triple negative pure DCIS	TN DCIS with invasive disease (n=10)	TN pure DCIS (n=9)
Triple negative pure DCIS compared to triple negative invasive breast disease	TN pure DCIS (n=9)	TN invasive breast disease (n=7)
HER2+ pure DCIS compared to ER+ pure DCIS	HER2 positive pure DCIS (n=7)	ER+ pure DCIS (n=8)
HER2+ pure DCIS compared to TN pure DCIS and ER+ pure DCIS	HER2 positive pure DCIS (n=7)	TN pure DCIS, ER+ pure DCIS (n=17)

Table 31: MIP Array Comparison Groups

[TN = triple negative; ER+ = oestrogen receptor positive]

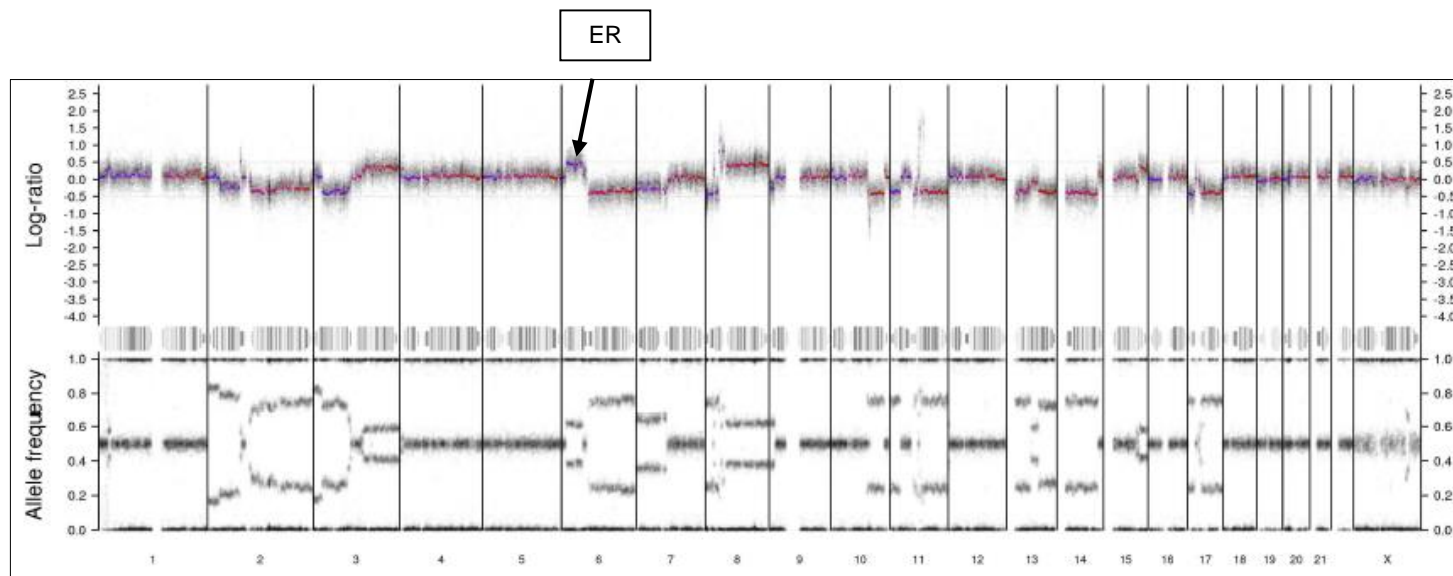
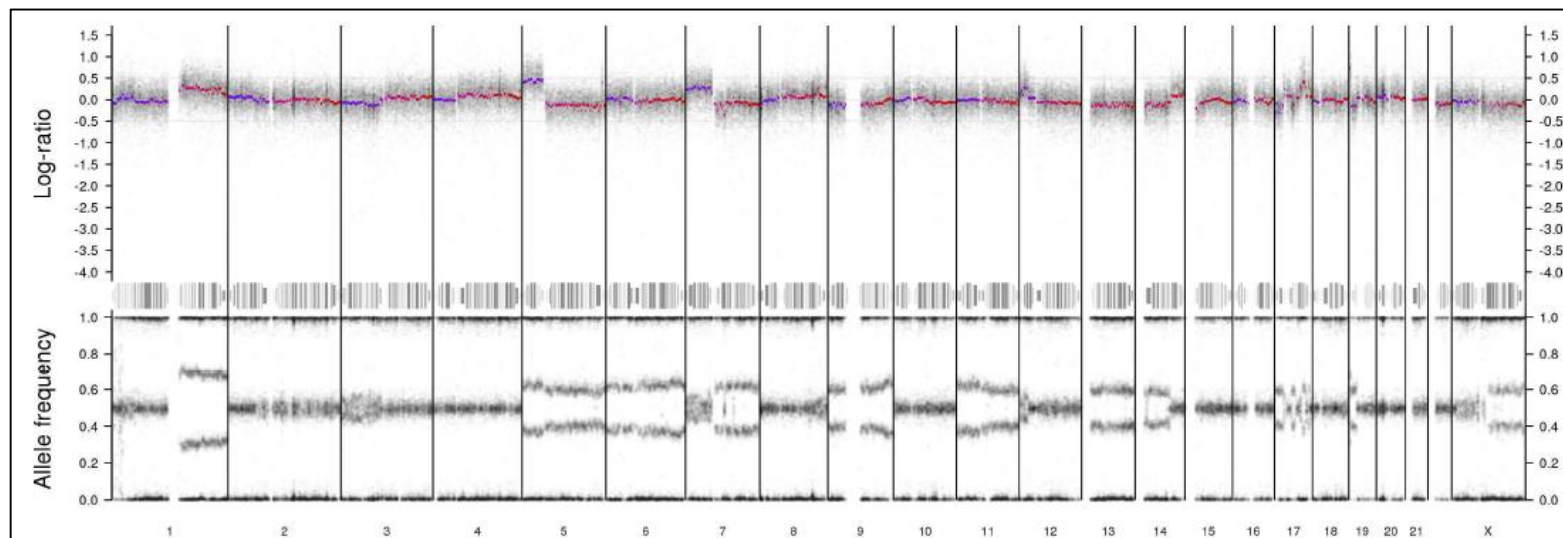


Figure 31: Chromosomal map for a ER Positive pure DCIS Case (DCISJPB10K 066).

The X axis in the diagram refers to chromosome number (1-21 and X) and the Y axis is split into two sections. The upper section shows log of ratio of signal of the same probe from the tumour and normal samples. The lower section of the plot shows allelic-ratio. Usually, this represents the ratio of B-allele as compared to A-allele. The value always lies between 0 and 1. e.g. if there is one copy each of both allele then allele ratio would be $1 / 2 = 0.5$ which represent the normal allele pairing. Each probe represents the number of copies for a genomic region. The probe signal from the normal sample is subtracted from the tumour samples and a log ratio (usually base 2)



Figure

Figure 32: Chromosomal Map for a Pure DCIS Triple Negative Case (DCISJPB10K001)

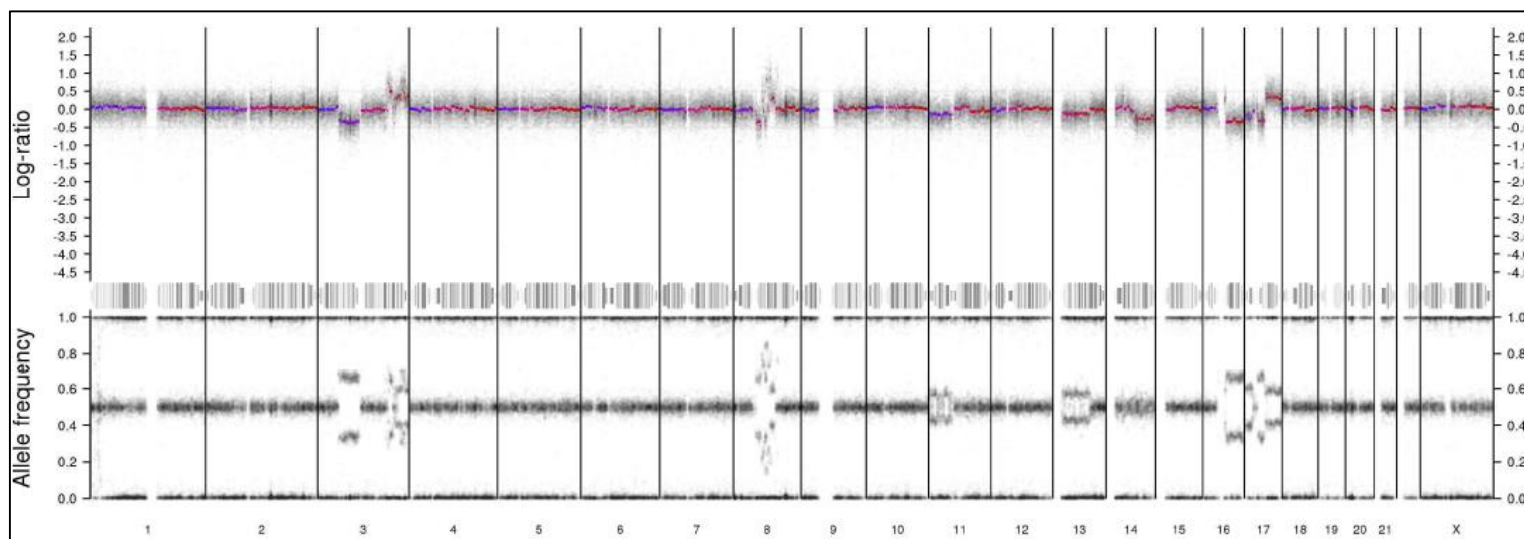


Figure 33: Chromosomal Map for a Pure HER2 Positive Case (DCISJPB10K070); note chromosome 17 allelic amplification.

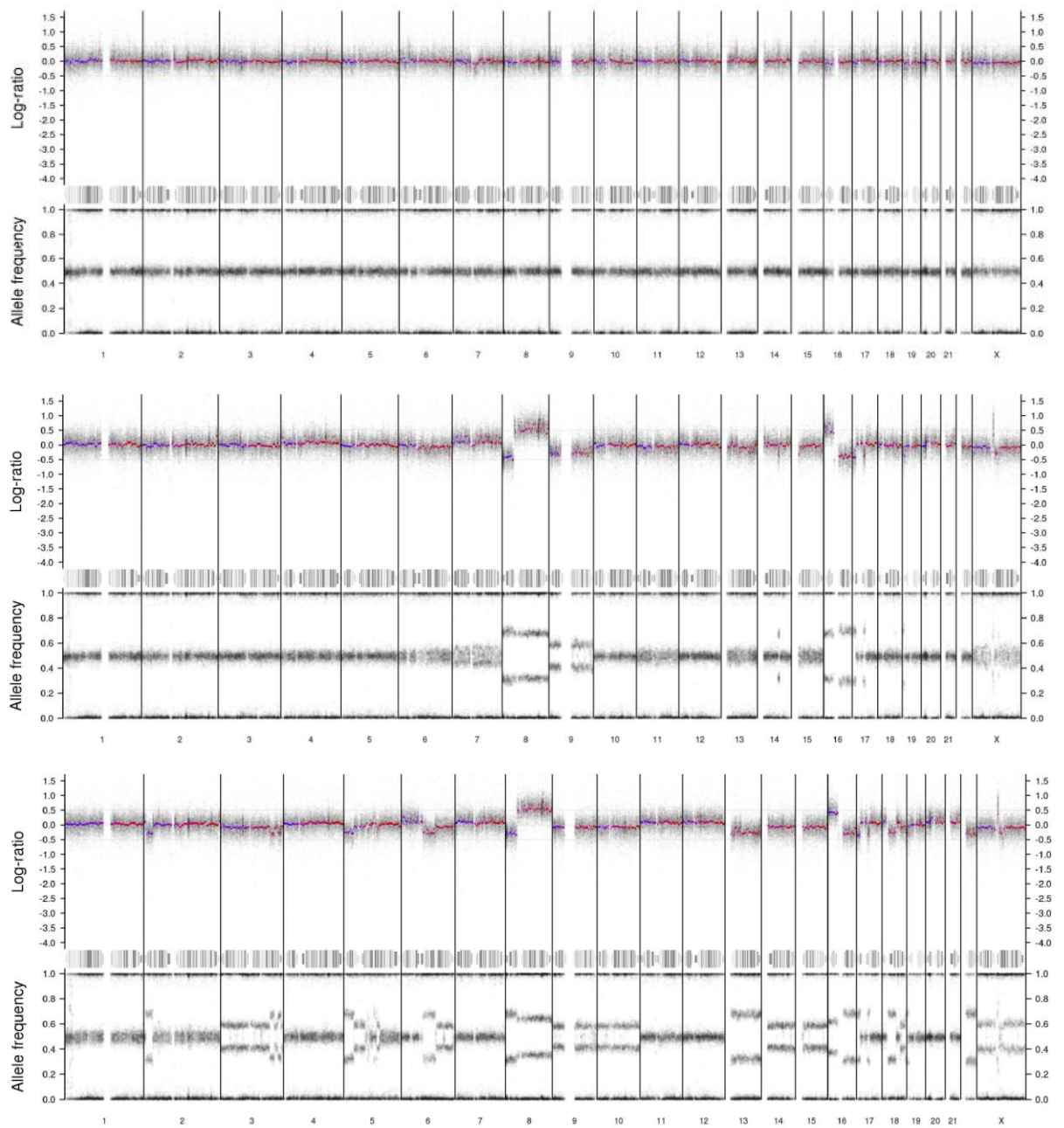


Figure 34: Chromosomal map for normal, DCIS and invasive samples from an ER positive lesion.

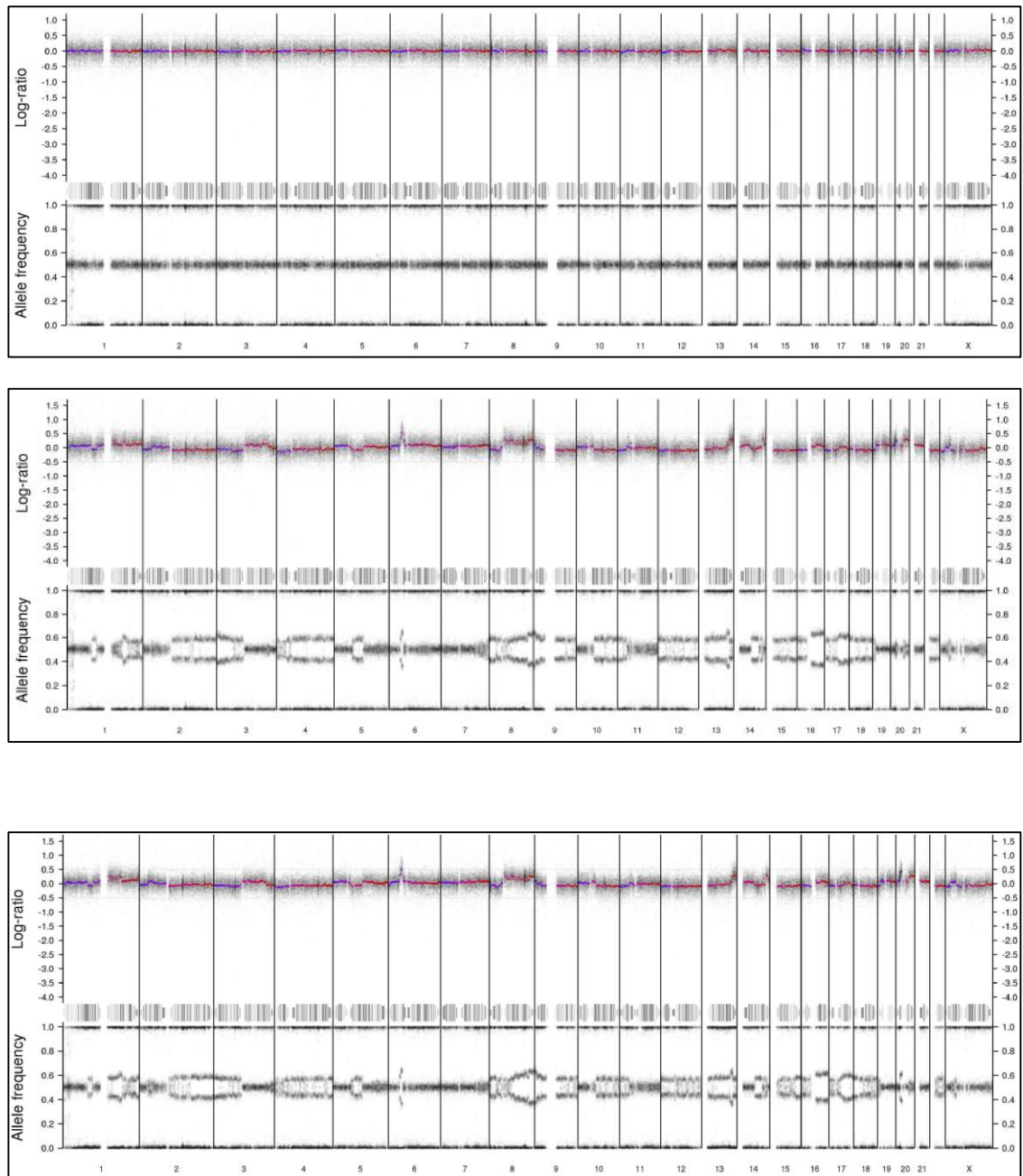


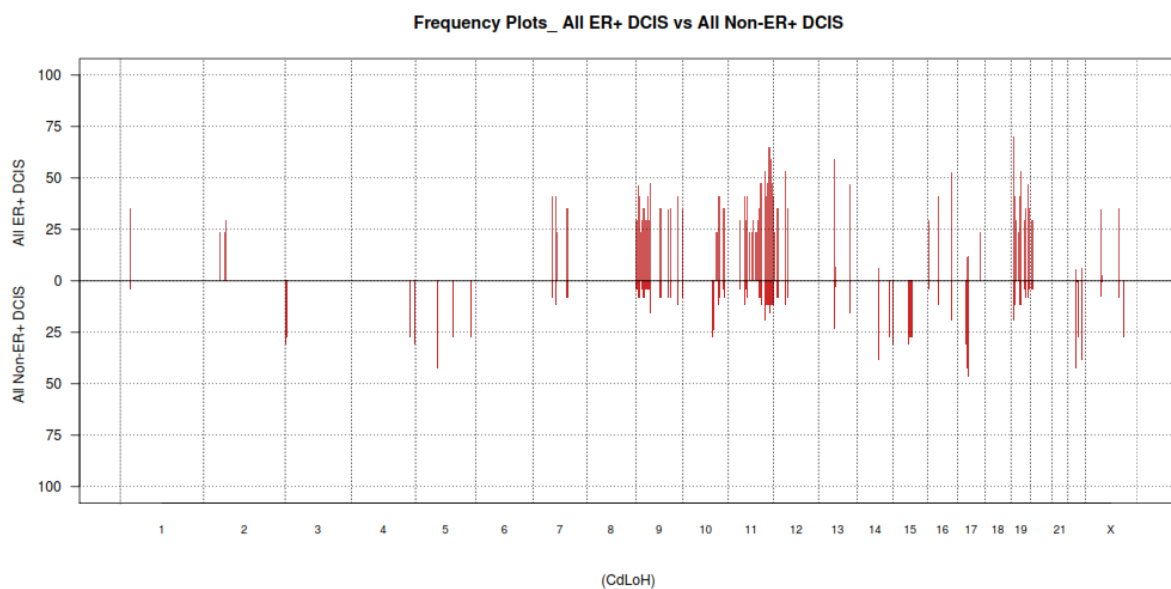
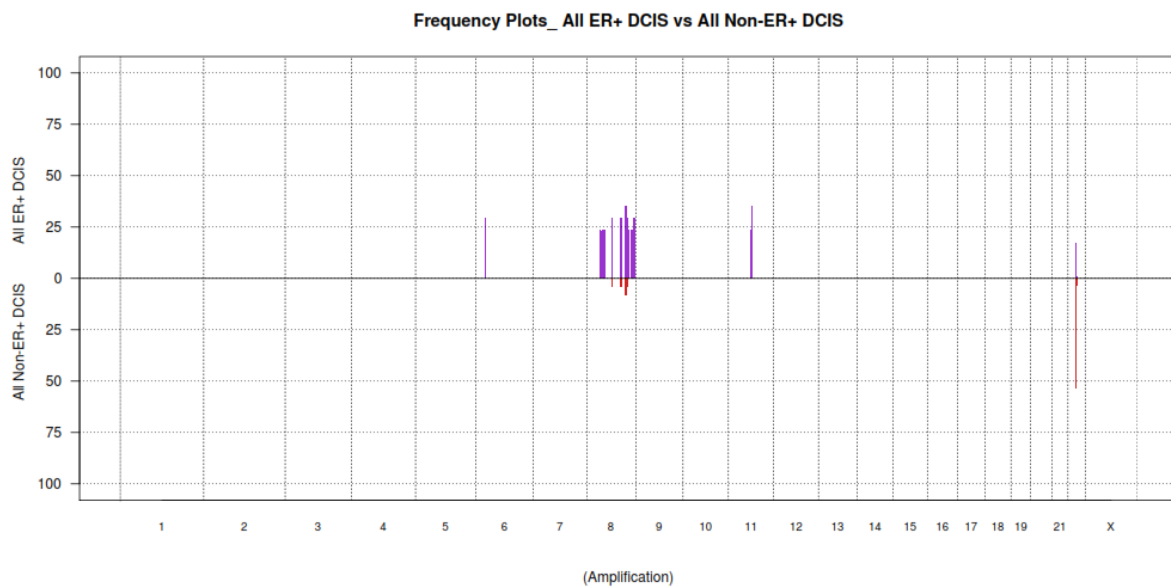
Figure 35: Chromosomal Map for normal, DCIS and invasive samples from a triple negative case.

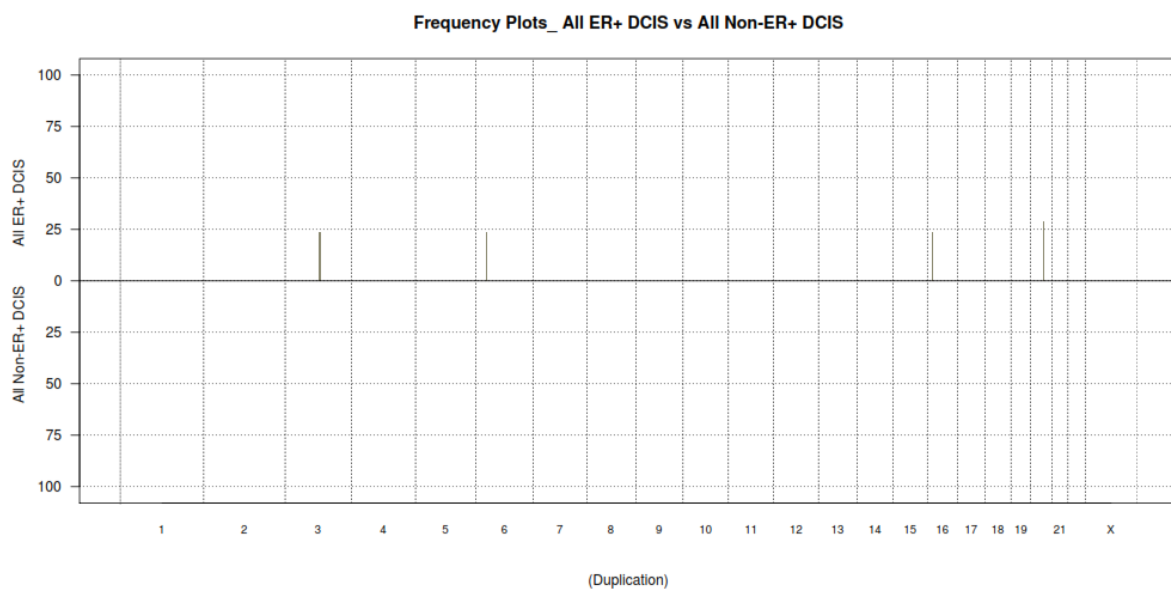
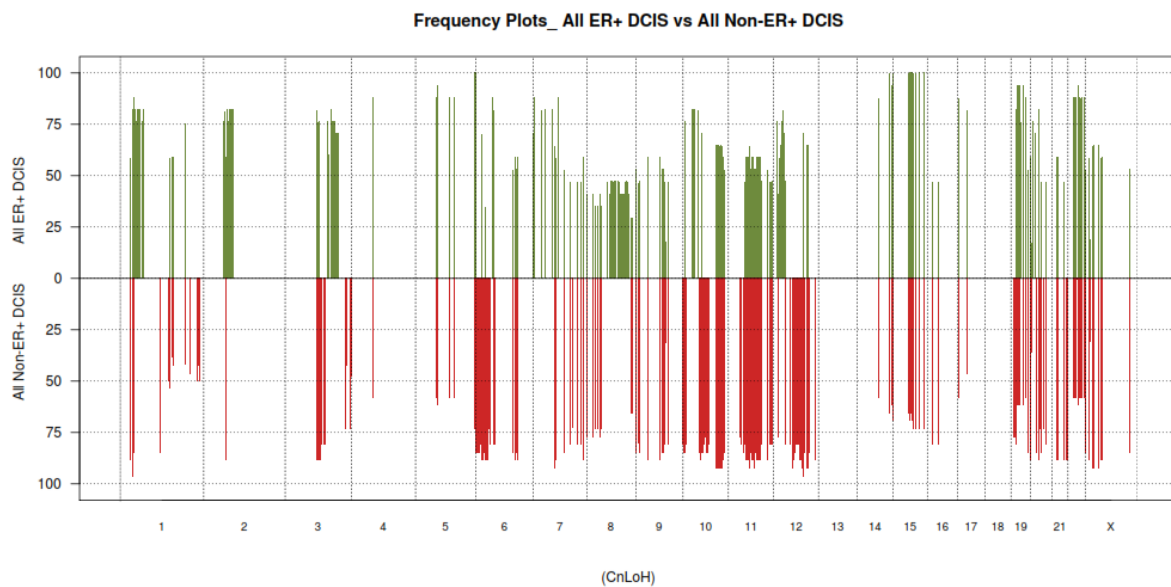
5.8.3 Copy Number Aberrations for Oestrogen Receptor Positive DCIS Compared to Non Oestrogen Receptor DCIS

These series examines the difference between oestrogen receptor positive DCIS and non oestrogen receptor positive DCIS.

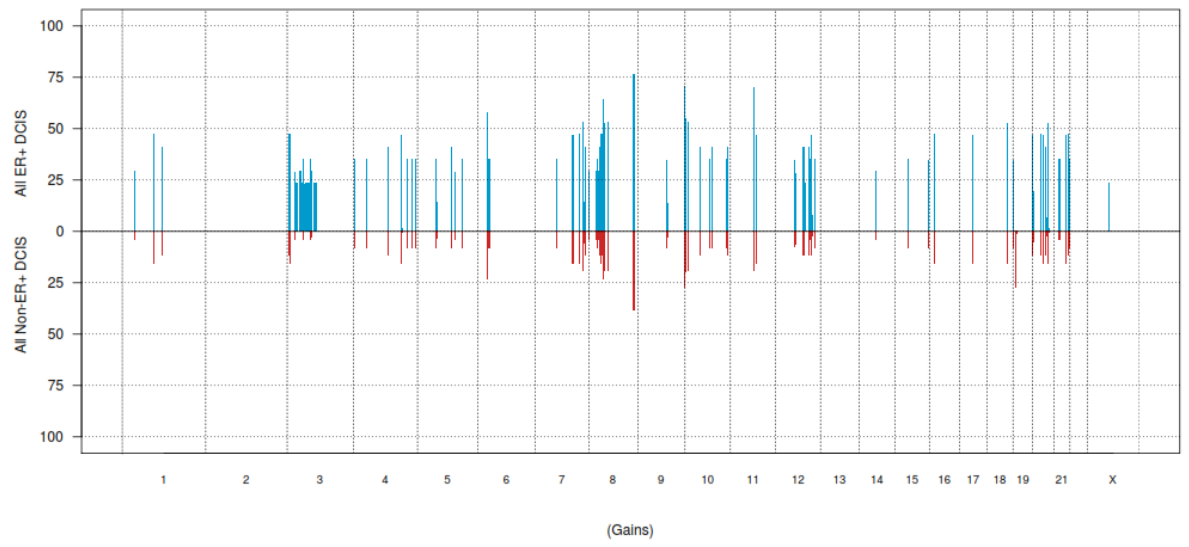
Frequency plots showing copy number aberrations between Oestrogen receptor positive DCIS and oestrogen receptor negative DCIS were provided by Breakthrough Breast Cancer/Research Oncology, King's College London Bioinformatics Department (Figure 37).

Figure 36: Frequency plots showing copy number aberrations between oestrogen receptor positive DCIS and non oestrogen positive DCIS (pages 165-167).

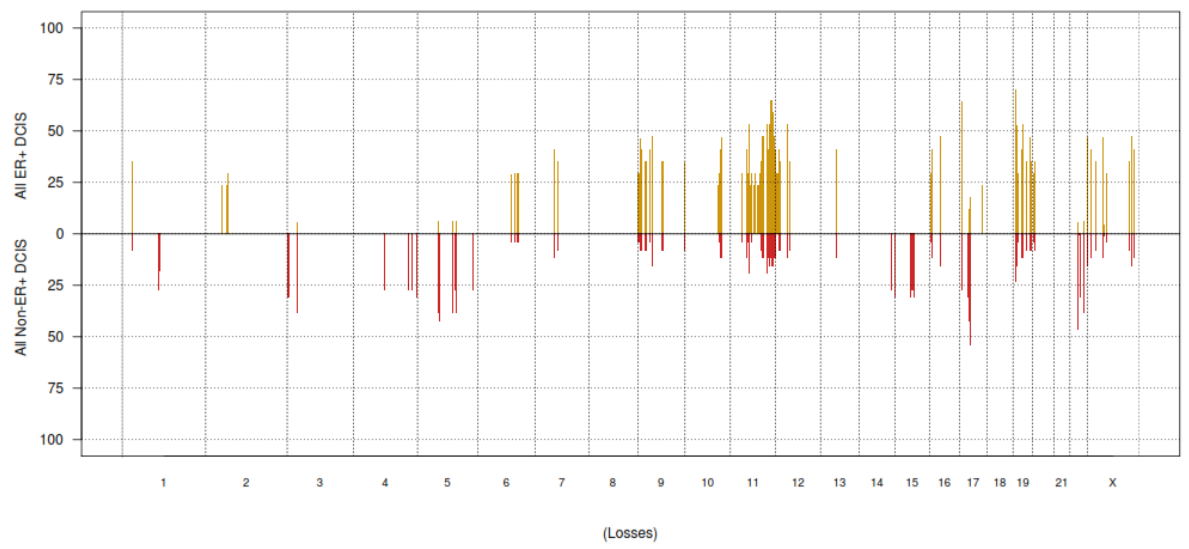


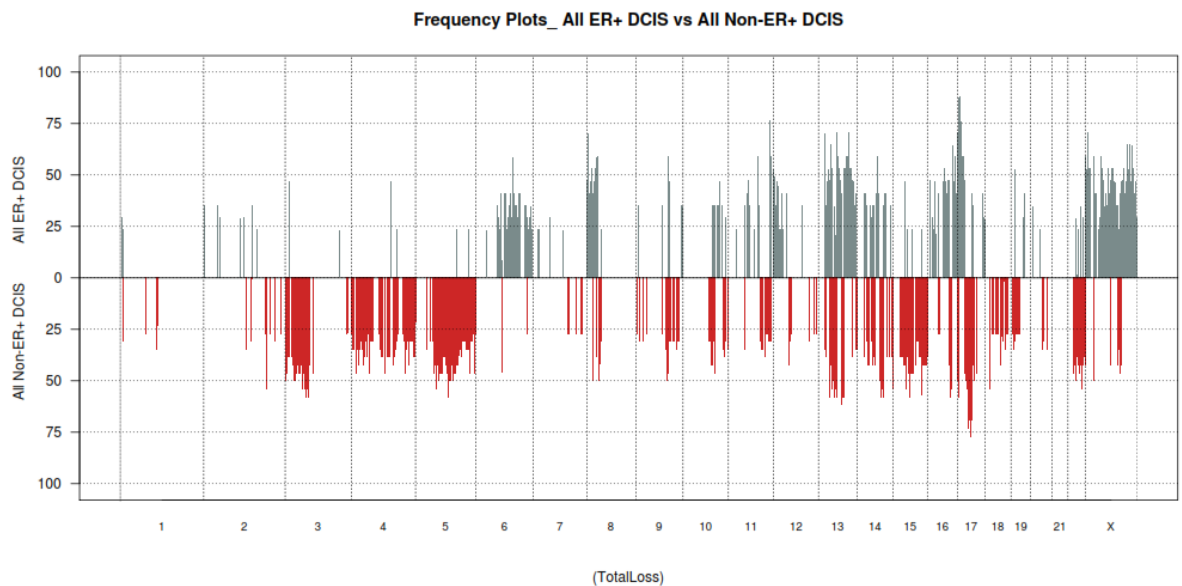
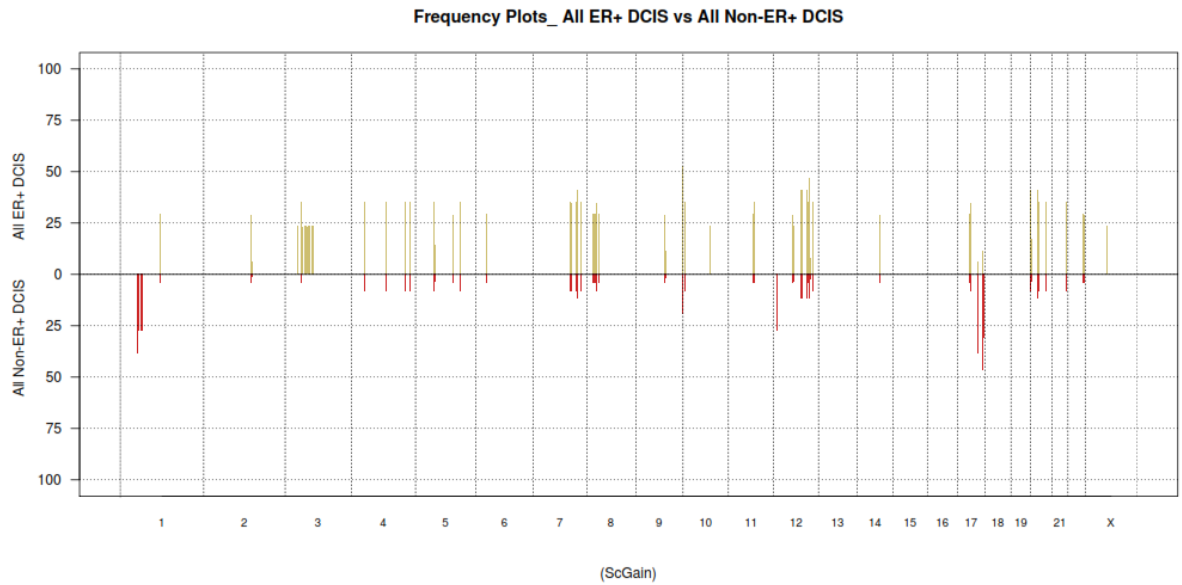


Frequency Plots_ All ER+ DCIS vs All Non-ER+ DCIS



Frequency Plots_ All ER+ DCIS vs All Non-ER+ DCIS





5.8.3.1 Amplification in Oestrogen Receptor (ER) Positive DCIS and Oestrogen Receptor Negative DCIS

There are amplifications found in ER positive DCIS (n=5/17) not observed in the majority of non ER DCIS (n=2/26).

1. p-values < 0.05;
2. A gene list is mapped from 16 genomic regions found on chromosomes 8, 11;
3. These regions encompass 118 genes altered in ER positive DCIS;

Genes:

5S_rRNA	AC011626.1	AC021636.1	AC027238.1	AC090821.2	AC090821.3	AC090821.4	AC090821.5	AC103706.1	AC103764.1
	AC103863.1	AC107374.1	AC115836.1	AF130342.1	AF178030.2	AF186191.1	AF230666.2	AL031777.1	ANKRD46
	AP003716.2	AP003716.4	AP003716.5	CCND1	CEBPD	CHAC1	CSMD3	CTC-458A3.1	CTD-2182N23.1
	DENND3	DEPDC6	EIF2C2	FGF19	FGF3	FGF4	GDAP1	HFE	HIST1H1C
	HIST1H1E	HIST1H1PS1		HIST1H1T	HIST1H2AB	HIST1H2AC	HIST1H2AD	HIST1H2AE	HIST1H2APS3
	HIST1H2APS4		HIST1H2BB	HIST1H2BC	HIST1H2BD	HIST1H2BE	HIST1H2BF	HIST1H2BG	HIST1H2BH
	HIST1H3A	HIST1H3B	HIST1H3C	HIST1H3D	HIST1H3E	HIST1H3F	HIST1H3G	HIST1H4A	HIST1H4B
	HIST1H4D	HIST1H4E	HIST1H4F	HIST1H4G	hsa-mir-151	hsa-mir-30b	hsa-mir-30d	IKBKB	KB-1615E4.1
	KIAA0146	KIAA0196	MYEOV	NCALD	NDRG1	NSMCE2	OAOV1	PHF20L1	POLB
	PXDNL	RAD21	RNF19A	RP11-1023P17.1		RP11-231D20.1		RP11-697N18.1	RP11-709P2.1
	221C16.7	RPS10P1	RRM2B	SLA	SLC45A4	SNORA40	SNORA7	SPAG1	SQLE
	ST3GAL1	TG	TMEM71	TRAPPC9	TRPS1	U6	U7	U91328.2	UNC93B5
	Y_RNA	ZFAT	ZFATAS	ZMAT4					WISP1

PANTHER analysis: 56 mapped ids are found, 62 mapped ids are not found.

There are amplifications found in non ER DCIS (n=2/26) that are also present in some ER positive DCIS (n=6/17).

1. p=values <0.05;
2. A gene list is mapped from 13 genomic regions on chromosome 8;
3. These regions encompass 49 genes altered in non ER DCIS;

Genes:

AC021636.1	AC090821.2	AC090821.3	AC090821.4	AC090821.5	AC103706.1	AC103764.1	AC107374.1	AC115836.1	AF186191.1
AF230666.2	CEBPD	CSMD3	CTC-458A3.1		CTD-2182N23.1	DENND3	DEPDC6	hsa-mir-30b	hsa-mir-30d
	IKBKB	KCNK9	KIAA0146	KIAA0196	NDRG1	NSMCE2	PHF20L1	POLB	PRKDC
	PTK2	PXDNL	RP11-1023P17.1		RP11-231D20.1		RP11-697N18.1	RP11-709P2.1	SLA
	SLC45A4	SNORA40	SNORA7	SQLE	ST18	ST3GAL1	TG	TMEM71	TRAPPC9
	WISP1	ZFAT	ZFATAS	ZMAT4					

PANTHER analysis: 25 mapped ids are found, 24 mapped ids are not found.

5.8.3.2 Duplication in Oestrogen Receptor Positive DCIS and Oestrogen Receptor Negative DCIS

There are no duplications in ER negative DCIS (check the status of the HER2 positives) in this series.

There are duplications in ER positive DCIS (n=4/17) not observed in non ER positive DCIS (n=0/26)

1. p-values < 0.05;
2. A gene list is mapped from 4 genomic regions found on chromosomes 3, 6, 16;
3. These regions encompass 27 genes altered in ER positive DCIS;

Genes:

AC136443.1	AC138932.2	AL645941.2	AL669918.1	BFAR	HLA-DMB	HLA-DOB	HLA-DQA2	HLA-DQB2	HLA-DQB3	HLA-
Z	hsa-mir-1972	NIT2	NOMO1	NPIP	PDXDC1	PLA2G10	PPP1R2P1	PSMB8	PSMB9	TAP1
	XXbac-BPG181M17.5		XXbac-BPG246D15.8		XXbac-BPG246D15.9	Y_RNA				

PANTHER analysis: 15 mapped ids are found, 12 mapped ids are not found.

5.8.3.3 Genomic Gains in Oestrogen Receptor Positive DCIS and Non Oestrogen Receptor Positive DCIS

There are genomic gains in ER positive DCIS (n=6/17) not observed in non ER positive DCIS (n=0/26).

1. p-values < 0.05;
2. A gene list is mapped from 10 genomic regions on chromosome 3;
3. These regions encompass 75 genes altered in ER positive DCIS;

Genes:

5S_rRNA	AC096971.2		AC097359.1	AC097359.2	AC097359.3	AC099668.5	AC104186.1	AC104448.1	AC121252.1	AC121252.2
	AC126118.1	AC126118.2	AC141002.1	AMIGO3	APEH	ATRIP	C3orf35	C3orf54	CAMP	CCDC51
	CDC25A	CDH29	CELSR3	COL7A1	CTD-3211M3.1	FBXW12	FOXP1	GMPPB	GOLGA4	hsa-mir-2115
mir-711	IP6K1	IP6K2	LRRFIP2	MLH1	MST1	NCKIPSD	NDUFB1P1	NME6	PFKFB4	PLXNB1
	PRKAR2A	RNF123	RP11-20O23.1		RP11-24C3.2		RP11-391M1.3		RP11-430J3.1	
	RP11-430J3.2		RP11-502L5.2		RP11-520A21.1		RP11-528N21.1		RP11-528N21.2	
	RP11-528N21.3		RP11-905F6.1		RP13-1056D16.2	RP13-480C15.1	RPL14	SHISA5	SLC26A6	
snoU13	SPINK8	TMEM89	TREX1	U5	U6	U7	UBA7	UCN2	UQCRC1	Y_RNA
	ZNF589	ZNF619	ZNF620	ZNF621						

PANTHER analysis: 36 mapped ids are found, 39 mapped ids are not found.

There were gains present in non ER DCIS (n=10/17) which are also present in ER positive DCIS (n=13/10).

1. p-values < 0.05;
2. A gene list is mapped from 16 genomic regions found on chromosomes 6, 8, 10, 19;

- These regions encompass 30 genes altered in both non ER DCIS and ER DCIS;

Genes:

5S_rRNA	AC008734.2	AC090821.2	AC090821.3	AC090821.4	AC090821.5	AC103706.1	AF186191.1	AF216667.1	AF230666.2	AL713922.2
	CTC-458A3.1		CTD-2182N23.1		DIP2C	IKBKB	IL9RP2	LRRC6	MUC16	NDRG1
	PHF20L1	RP11-631M21.1		RP11-631M21.2		RP11-631M21.6		SLA	SNORA40	ST3GAL1
	TMEM71	WISP1	ZMYND11							TG

PANTHER analysis: 12 mapped ids are found, 18 mapped ids are not found.

5.8.3.4 Genomic Sc Gains in Oestrogen Receptor Positive DCIS and Non Oestrogen Positive Receptor DCIS

There are some overlaps in Sc gains found in ER positive DCIS (n=9/17) and non ER positive DCIS (n=4/26).

- p-values < 0.05;
- A gene list is mapped from 49 genomic regions found on chromosomes 3, 4, 5, 7, 8, 10, 11, 12, 17, 20, 21;
- These regions encompass 180 genes altered in for ER positive DCIS and some non ER positive DCIS;

Genes:

5S_rRNA	ABCG1	ABHD12	AC004001.1	AC005355.1	AC008038.1	AC008264.4	AC008937.2	AC009365.4	AC009518.8	AC013434.1
	AC016831.7	AC018642.1	AC058791.1	AC058791.2	AC079595.1	AC079595.2	AC080112.3	AC087071.1	AC087071.2	AC092214.10
	AC097534.1	AC097534.2	AC105285.1	AC105285.2	AC121247.2	AC121252.2	AC135506.1	AC137630.1	AF213884.1	AL031666.1
AL034548.1	AL034548.2	AL034548.3	AL109954.2	AL121894.1	AL157718.1	AL365356.1	AL713922.2	AL732437.1	ANKS1B	
AP002827.1	ARIH2	ARL1	ASB13	C10orf18	C12orf23	C20orf26	C21orf121	C21orf128	C2CD2	C3orf62
C3orf71	C5orf15	CALML3	CALML5	CCDC25	CCDC36	CCDC71	CDC6	CDKL3	CDKN2AIPNL	
CHCHD3	CRNKL1	CRY1	CST11	CST3	CST8	CST9	CST9L	CSTL1	CTD-2410N18.1	
DALRD3	DIP2C	DOCK4	EIF4E2P1	ENTPD6	EPHA1	GALNT7	GDPD4	GPX1	GZF1	
HMGB2	hsa-mir-191	hsa-mir-425	hsa-mir-620	IL9RP2	IMPDH2	KLHDC10	KLHDC8B	LAMB2	MAP3K1	
MED13L	MKLN1	MTERFD3	NAPB	NAT5	NDUFAF3	NET1	NFKB1	NRSN2	NXT1	
P4HTM	PAK1	PODXL	PPIAP2	PPP2CA	PRKAR2A	PYGB	QARSQRICH1	RALGAPA2	RARA	
RHOA	RHOH	RIN2	RP11-10K16.1		RP11-135F9.1		RP11-215P8.1		RP11-215P8.2	
	RP11-215P8.3		RP11-215P8.4		RP11-218C14.2	RP11-218C14.6	RP11-218C14.7		RP11-23O13.1	
	RP11-306G20.1		RP11-312A15.2		RP11-312A15.3		RP11-316M24.1		RP11-336A10.4	
	RP11-336A10.5		RP11-336A10.7		RP11-395I6.2		RP11-3B7.1	RP11-3B7.6	RP11-148H17.1	
	RP11-631M21.1		RP11-631M21.2	RP11-631M21.6		RP11-694I15.6		RP11-775D22.2	RP11-798M19.3	
	RP13-131K19.1	RP13-131K19.2		RP13-463N16.6		RP3-322G13.5		RP3-333B15.2	RP4-569M23.2	
	RP5-1002M8.4	RP5-1103G7.4		RPS27P23	SAP30	SKP1	SLC24A3	SLC25A20	snoU13	SOX12
									SPIC	

TCF7	TMTC2	TOP2A	TRIB3	U5	U6	UBE2B	UBE2H	UMODL1	USP19	USP4
UTP20	VDAC1	WDR6	Y_RNA	ZC3HC1	ZCCHC3	ZMYND11	ZMYND8	ZNF295		

PANTHER analysis: 87 mapped ids are found, 92 mapped ids are not found.

There are some overlaps in Sc gains for non ER positive DCIS (n=12/26) and ER positive DCIS (n=9/17).

1. p-values < 0.05;
2. A gene list is mapped from 25 genomic regions found on chromosomes 1, 7, 10, 12, 17.
3. These regions encompass 83 genes altered in non ER positive DCIS and ER positive DCIS.

Genes:

5S_rRNA	AC013434.1	AC016182.1	AC018628.1	AC018628.3	AC058791.2	AC079595.1	AC079595.2	AC096947.1	AGBL4		
	AL713922.2	ANKS1B	ARMC7	BRIP1	C12orf23	C1orf123	C1orf185	C20orf26	CDKN2C		
	CFLP2	CPT2	CRNKL1	CRY1	CSDA	CYTH1	DIP2C	DMRTA2	DNAH17	FAF1	HN1
	IL9RP2	INTS2	LRP8	MAGOH	MED13	MKLN1	MTERFD3	NAT5	NFIA		
	NT5C	PRH1	PRH2	PRR4	RIN2	ROR1	RP11-117D22.1		RP11-183G22.2		
	RP11-183G22.3		RP11-24J23.2		RP11-323N12.1		RP11-323N12.2		RP11-631M21.1		
	RP11-631M21.2		RP11-631M21.6		RP4-784A16.1	RP4-784A16.2		RP4-784A16.3	RP4-		
784A16.4	RP4-784A16.5		RP5-1002M8.4		RP5-1024G6.2		RP5-1024G6.5		RP5-833A20.1		
	RP5-850O15.3		RP5-926E3.1		SLC16A5	TAS2R10	TAS2R13	TAS2R14		TAS2R19	
	TAS2R20	TAS2R31	TAS2R50	TAS2R7	TAS2R8	TAS2R9	TIMP2	TMTC2	U6	USP36	
	Y_RNA	ZMYND11	ZNF859P								

PANTHER analysis: 46 mapped ids are found, 47 mapped ids are not found.

5.8.3.5 Losses in Oestrogen Receptor Positive DCIS and Non Oestrogen Receptor Positive DCIS

There are genomic losses found in ER positive DCIS (n=8/17) not observed in non ER positive DCIS (n=0/26).

1. p-values < 0.05;
2. A gene list is mapped from 25 genomic regions found on chromosomes 2, 7, 9, 10, 11, 1;
3. These regions encompass 79 genes altered in ER positive DCIS;

Genes:

AC006432.1	AC007403.1	AC007403.2	AC007537.1	AC007537.2	AC009474.1	AC009474.2	AC022882.1	AC036111.1	AC068339.1	
AC079920.1	ACSM4	AP000720.1	AP001482.1	AP002826.1	AP003065.3	AP006437.1	APLNLR	ARAP1	ARHGEF17	
BCL2L14	CACNA1C	CD163L1	ETV6	FAM168A	GRM5	LRP6	LRRC55	MANSC1	OR10AG1	OR5AK3P
OR5AK4P	OR5AQ1P	OR5AS1	OR5BE1P	OR5D13	OR5D14	OR5D16	OR5D18	OR5F1	OR5G3P	OR5G5P
OR5I1	OR5J1P	OR5J2	OR5L1	OR5L2	OR5T1	OR5T2	OR5T3	OR5W1P	OR5W2	
OR7E5P	OR8H1	OR8H2	OR8H3	OR8I2	OR8J3	OR8K5	OR8V1P	OR9M1P	P2RX3	
PRG2	PRG3	PTPRD	RELT		RP11-267J23.1		RP11-31L22.2	SBDS	SLC22A8	
SLC43A3	snoU13	SPRYD5	SSRP1	TNKS1BP1	TYR	TYW1	U2	U6		

PANTHER analysis: 46 mapped ids are found, 33 mapped ids are not found.

There are genomic losses found in non ER positive DCIS (n=11/26) not observed in ER positive DCIS (n=0/17).

1. p-values < 0.05;
2. A gene list is mapped from 17 genomic regions found on chromosomes 3, 4, 5, 14, 15, 17, 22;
3. These regions encompass 35 genes altered in non ER positive DCIS;

Genes:

AB019437.56	AC008772.1	AC008872.1	AC026202.1	AC026202.5	AC026214.2	AC092373.1	C22orf30	CD180	CNTN6	
FAT1	IGHV1-67	IGHV1-68	IGHV1-69	IGHV2-70	IGHV3-66	IGHV3-71	IGHV3-72	IGHV3-73	IGHV3-74	IGHVII-
67-1	IGHVIII-67-2	MAST4	PISD	RAB27A	RP11-173M1.1	RP11-287J9.1	RP11-287J9.2		RP11-308K2.1	
	RP11-5P22.1		RP11-5P22.2		RP11-5P22.3		RP5-858B16.5	SFRS12	snoU13	

PANTHER analysis: 6 mapped ids are found, 29 mapped ids are not found.

5.8.3.6 Total Losses in Oestrogen Receptor Positive DCIS and Non Oestrogen Receptor Positive DCIS

There were total losses present in ER positive DCIS (n=13/17) not observed in non ER positive DCIS (n=0/26).

1. p-values < 0.05;
2. A gene list is mapped from 11 genomic regions found on chromosomes 8, 11, 13, 17;
3. These regions encompass 88 genes altered in ER positive DCIS;

Genes:

5S_rRNA AC005696.2 AC005696.3 AC006435.1 AC015799.1 AC015799.2 AC027763.1 AC027763.2 AC032044.1 AC055839.2 AC087742.1
AC087742.2 AC090617.1 AC091153.4 AC116914.1 AC118754.2 AC118754.4 AC127521.1 AC130689.5 AIPL1 AL450226.1
AL450226.2 ALOX12 ALOX15 AP001970.1 ARRB2 ASGR1 ASGR2 ATP2A3 BCL6B BHLHA9
C17orf100 C17orf49 C17orf85 CAMKK1 CLEC10A CRKCSMD1 CTD-2545G14.1 DLG4 DPH1 FAM64A
FBXO39 GAS7 GGT6 GSG2 HIC1 hsa-mir-132 hsa-mir-195 hsa-mir-212 hsa-mir-497 ITGAE
KIAA0664 KIAA0753 MED31 METT10D MNT MYBBP1A OVCA2 P2RX1 PAFAH1B1 PELP1 PITPNM3
RNASEK RP11-314A20.1 RP11-314A20.2 RP11-609D21.1 RP11-76K19.5 SGSM2
SLC13A5 SLC16A11 SLC16A13 SMG6 SMTNL2 SNORD91 snoU13 SPNS2 SPNS3 SRR TEKT1
TSR1 TUSC5 TXNDC17 U6 U8 UBE2G1 XAF1 YWHAE

PANTHER analysis: 51 mapped ids are found, 37 mapped ids are not found.

There were total losses present in non ER positive DCIS (n= 20/26) not observed in ER positive DCIS (n=0/17).

1. p-values < 0.05;
2. A gene list is mapped from 21 genomic regions from chromosome 17;
3. These regions encompass 182 genes altered in non ER positive DCIS;

Genes:

5S_rRNA ABHD5 AC008734.2 AC011816.4 AC020626.1 AC025029.1 AC092038.1 AC092798.2 AC092798.3 AC093416.3
AC098477.1 AC098477.2 AC098647.1 AC099539.1 AC099544.1 AC099544.2 AC099557.1 AC099778.3 AC104163.1 AC104184.1
AC104307.1 AC104434.1 AC104434.2 AC104434.3 AC104445.1 AC105935.1 AC112512.1 AC115283.1 AC121251.1 AC121252.2
AC124914.3 AC133141.1 AC133680.1 AC135966.1 ANO10 C3orf35 C3orf39 CACNA1D CADM2 CAMKV
CDC25A CDCP1 CELSR3 CLEC3B CSPG5 CTNNB1 DHX30 EIF1B EIF4BP9 ENTPD3 EXOSC7
FAM116A FAM198A FOXP1 FYCO1 GBE1 hsa-mir-1226 hsa-mir-564 ITGA9 KIF15 LARS2
LIMD1 LRIG1 MAP4 MITF MLH1 MUC16 MYRIP NCKIPSD NGLY1 NKIRAS1
NR1D2 OXSM PRICKLE2 RARB RP11-1029M24.1 RP11-1029M24.2 RP11-107O13.1 RP11-129B22.1
RP11-129B22.2 RP11-129B22.4 RP11-129B22.5 RP11-136C24.1 RP11-136C24.2 RP11-141M3.2
RP11-142L1.2 RP11-142L1.3 RP11-148G20.1 RP11-188P20.2 RP11-260O18.1 RP11-26G10.1
RP11-272D20.2 RP11-30M20.1 RP11-341J3.1 RP11-353H3.1 RP11-372H2.1
RP11-391M1.2 RP11-391M1.3 RP11-395P16.1 RP11-420K5.1 RP11-564P9.1 RP11-681O4.1
RP11-697K23.1 RP11-755B10.3 RP11-814M22.1 RP11-88B8.2 RPL15 SACM1L
SLC6A20 SLMAP SMARCC1 SNORA64 SNORD5 SNRK TGM4 TMEM42 TOP2B TRAP U2
U3 U5 U6 U7 UBE2E1 UBE2E2 ULK4 Y_RNA ZDHHC3

PANTHER analysis: 50 mapped ids are found, 78 mapped ids are not found.

5.8.3.7 CdLOH in Oestrogen Receptor Positive DCIS and Non Oestrogen Receptor Positive DCIS

There is CdLOH present in ER positive DCIS (n=8/17) not observed in non ER positive DCIS (n= 0/26).

1. p-values < 0.05;
2. A gene list is mapped from 13 genomic regions found on chromosomes 9, 11;
3. These regions encompass 30 genes altered in ER positive DCIS;

Genes:

AC022882.1	AC036111.1	AP000866.1	AP000866.2	AP006437.1	C11orf61	ESAM	NRGN	OR5D13	OR5D14	OR5D16
OR5D18	OR5G3P	OR5G5P	OR5L1	OR5L2	OR5T1	OR5T2	OR5T3	OR8H1	OR8V1P	
P2RX3	PRG2	PRG3	PTPRD	snoU13	SSRP1	TNKS1BP1	U2	VSIG2		

PANTHER analysis: 19 mapped ids are found, 11 mapped ids are not found.

There is CdLOH in non ER positive DCIS (n=11/26) not observed in ER positive DCIS (n=0/17).

1. P-values < 0.05;
2. A gene list is mapped from 13 genomic regions found on chromosomes 8, 9, 10, 11;
3. These regions encompass 27 genes altered in non ER positive DCIS;

Genes:

AB019437.56	AC008772.1	AC026214.2	AC092373.1	CNTN6	FAT1	IGHV1-67	IGHV1-68	IGHV1-69	IGHV2-70
IGHV3-66	IGHV3-71	IGHV3-72	IGHV3-73	IGHV3-74	IGHVII-67-1	IGHVIII-67-2	MAST4	RP11-173M1.1	
RP11-287J9.1		RP11-287J9.2		RP11-308K2.1		RP11-5P22.1		RP11-5P22.2	
RP11-5P22.3		SFRS12	snoU13						

PANTHER analysis: 3 mapped ids are found, 24 mapped ids are not found.

5.8.3.8 CnLOH in Oestrogen Receptor Positive DCIS and Non Oestrogen Receptor Positive DCIS

There is CnLOH in ER positive DCIS (n=15/17) not observed in non ER positive DCIS (n=0/26).

1. p-values < 0.05;

2. A gene list is mapped from 40 genomic regions found on chromosomes 1, 2, 3, 6, 7, 10, 12 19, 20;
3. These regions encompass 200 genes altered in ER positive DCIS;

Genes:

5S_rRNA	7SK	AC004906.3	AC007040.5	AC008686.2	AC010733.4	AC012671.2	AC022021.2	AC062029.1	AC074389.1	
		AC074389.6	AC074389.7	AC074389.9	AC078841.5	AC083864.3	AC083864.4	AC104134.2	AC104170.1	AC110781.3
										AC110781.5
ACOT11	AKIRIN1	AL031289.1	AL031985.1	AL139244.1	AL160287.2	AL161740.1	AL161932.2	AL929472.1	AL929472.3	
APBB1IP	ARHGAP12	BAZ1B	C10orf50	C1orf122	C1orf168	C1orf175	C1orf191	C1orf228	C2orf51	C8A
C8B	CARD11	CDCP2	CITED4	CST1	CST4	CSTP2	CTPS	CYB5RL	DAB1	DEM1
DEPDC1	EDN2	EEPDC1	EGFR	EIF2AK3	ELFN1	EPHA10	EPS15	FAM151A	FHL3	FOXI3
FOXO6	GLIS1	GLUDP5	GRIK3	HEYL	HIVEP3	HPCAL4	hsa-mir-23a	hsa-mir-24-2	hsa-mir-27a	hsa-mir-30c-1
hsa-mir-30e	IL20RB	INPP5B	KCNQ4	KIF2C	KIF5B	MAD1L1	MANEAL	MPP7	MRPL37	MTF1
NCK1	NDUFS5	NFYC	NRP1	NT5C1A	OMA1	OSBPL3	OSBPL9	OVCH1	PABPC4	PARS2
PPAP2B	PPIE	PRKAA2	RET	RIMS3	RNF220	RP11-109P14.1			RP11-109P14.10	PDSS1
RP11-109P14.2		RP11-109P14.8		RP11-109P14.9		RP11-128B16.3			RP11-15J6.1	RP11-167O6.2
RP11-191G24.1	RP11-191G24.2		RP11-213P13.1		RP11-240D10.2		RP11-240D10.4			RP11-241I20.1
RP11-241I20.3		RP11-241I20.4		RP11-241I20.5		RP11-253A20.1			RP11-269F19.2	
RP11-275F13.1		RP11-275F13.3		RP11-342D11.2		RP11-348A7.1			RP11-348A7.2	
RP11-351M16.1		RP11-351M16.2		RP11-351M16.3		RP11-377K22.2			RP11-377K22.3	
RP11-378I13.1		RP11-399E6.1		RP11-399E6.4		RP1-144F13.3			RP11-472N13.3	RP11-
486B10.3	RP11-486B10.4		RP11-524K22.1		RP11-575C1.1	RP11-656D10.3			RP11-656D10.5	
RP11-656D10.6		RP11-781D11.1		RP11-85F14.1		RP11-85F14.6			RP1-28O17.1	
RP13-16H11.1		RP13-16H11.2	RP13-16H11.5		RP13-16H11.6		RP13-16H11.7		RP1-63P18.2	
RP4-614N24.1	RP4-678E16.4		RP4-694A7.2		RP4-694A7.3		RP4-694A7.4		RP4-705F19.1	
RP4-705F19.2		RP4-710M16.2		RP4-739H11.1		RP4-739H11.3		RP4-783C10.3		RP5-
1033K19.2	RP5-1066H13.4		RP5-1103B4.3		RP5-866L20.1		RP5-866L20.2		RP5-866L20.3	
RP5-997D24.3		RP5-997D24.5	RP6-239D12.1		RPE65	RPIA	SDK1	SF3A3	SNORA63	
SNORD112	snoU13	SSBP3	TCEB1P18		TMEM53	TMTC1	TTC39A	TTC4	U3	U5
U7	UTP11L	VAX2	WAC	Y_RNA	YRDC	ZNF642	ZNF643	ZNF684	ZSWIM4	U6

PANTHER analysis: 77 mapped ids are found, 123 mapped ids are not found.

There is CnLOH present in non ER positive DCIS (n=24/26) not observed in ER positive DCIS (n=0/17).

1. p-values < 0.05;
2. A gene list is mapped from 38 genomic regions found on chromosomes 1, 3, 6, 10, 12, 21;
3. These regions encompass 126 genes altered in non ER positive DCIS;

Genes:

5S_rRNA	7SK	AACS	ABI3BP	AC078811.1	AC083805.1	AC083805.2	AC106728.2	AC108739.1	AC117377.1	
ADPRHL2	AF015262.1	AF015262.2	AL008732.1	AL034372.1	AL136180.1	AL136230.1	AL138724.3	AL138889.1	AL139044.1	
AL139286.1	AL590403.1	AL773603.1	AL844908.5	ANKS1B	AP000304.12	AP000313.1	AP000313.2	AP000318.2	AP000320.6	
AP000322.53	AP000322.54		AP000569.2	AP000569.8	AP001630.1	AP001630.5	AP001631.9	ARHGAP9	ATP5O	
C21orf51	C21orf67	C21orf70	C6orf125	CAP2	CBS	CLDND1	CLIC6	COL8A2	CPOX	CTD-2021J15.1
CTD-2021J15.2	DCBLD2	DCTN2	EIF2C1	EIF2C3	EIF2C4	FAM8A1	GNL1	GPR15	GRM4	hsa-mir-1275
mir-548a-1	IFNGR2	IMPG2	IP6K3	ITGB2	ITPR3	ITSN1	KCNE1	KCNE2	KIF13A	KIF5A
LEMD2	MARS	MGAT4C	MIP	MLN	MRPL51P2	NCRNA00160	NDUFV3	NUP153	OR5K2	PKNOX1
PRKG1	PRR3	PTTG1IP	RCAN1	RNF144B	RP11-14H3.3		RP11-204E9.1		RP11-227H4.1	
RP11-227H4.4		RP11-227H4.5	RP11-254A17.1		RP11-301G23.1		RP11-319J24.1		RP11-319J24.3	
RP11-377N20.1	RP11-40M23.1		RP11-524C21.1		RP11-569H14.1		RP11-663C11.1		RP11-686D16.1	
RP11-686D16.2		RP1-273P12.3		RP1-67M12.1		RP3-322L4.2	RP3-468B3.2		RP4-665N4.4	
RP4-665N4.5		RUNX1	SNORA11	snoU13	SOX4	ST3GAL6	SYT1	TEKT2	TIMELESS	
TMEM50B	TMTC2	TRIM38	U2AF1	U6	U7	Y_RNA				

PANTHER analysis: 56 mapped ids are found, 70 mapped ids are not found.

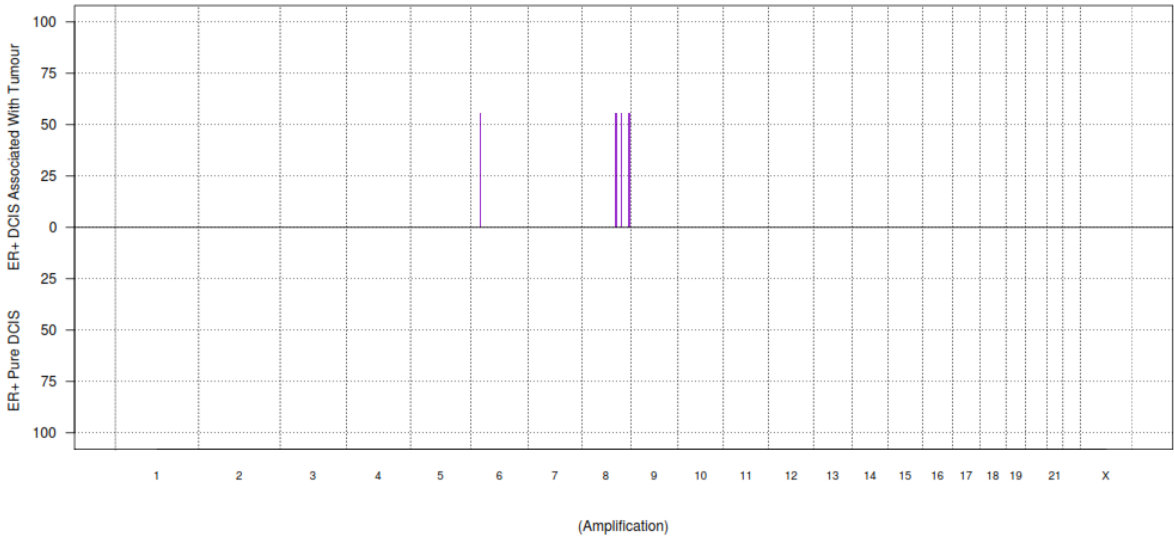
5.8.4 Copy Number Aberrations for Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease

These series examines the difference between oestrogen receptor positive pure DCIS and oestrogen receptor positive DCIS associated with invasive breast disease i.e. the DCIS component and not the invasive breast disease.

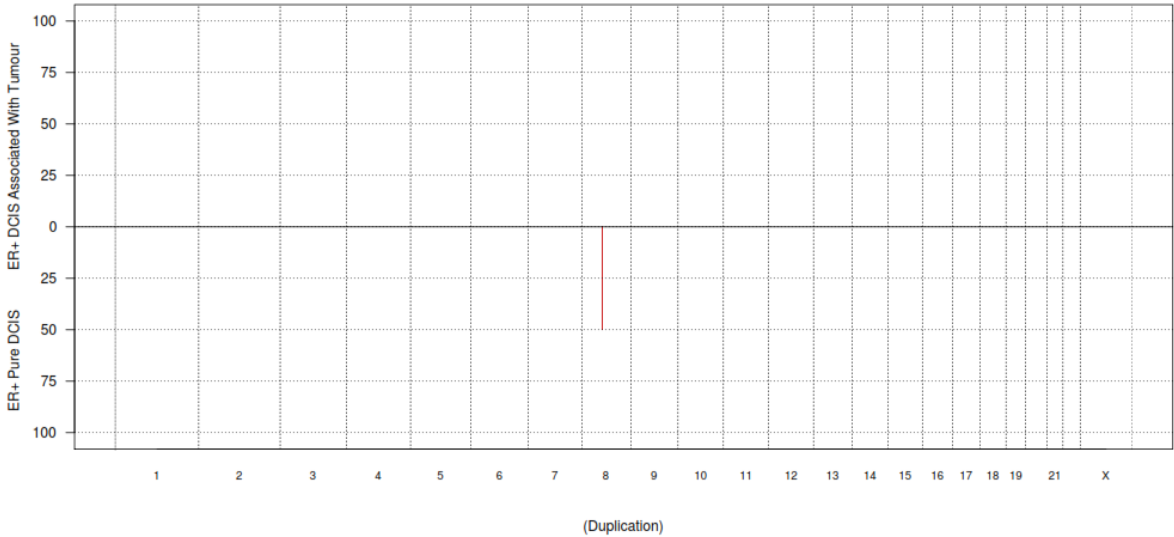
Frequency plots showing copy number aberrations between oestrogen receptor positive pure DCIS and oestrogen receptor positive DCIS associated with invasive breast disease were provided by Breakthrough Breast Cancer/Research Oncology, King's College London Bioinformatics Department (Figure 37).

Figure 37: Frequency plots showing copy number aberrations between oestrogen positive pure DCIS and oestrogen positive DCIS associated with tumour (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 177-180).

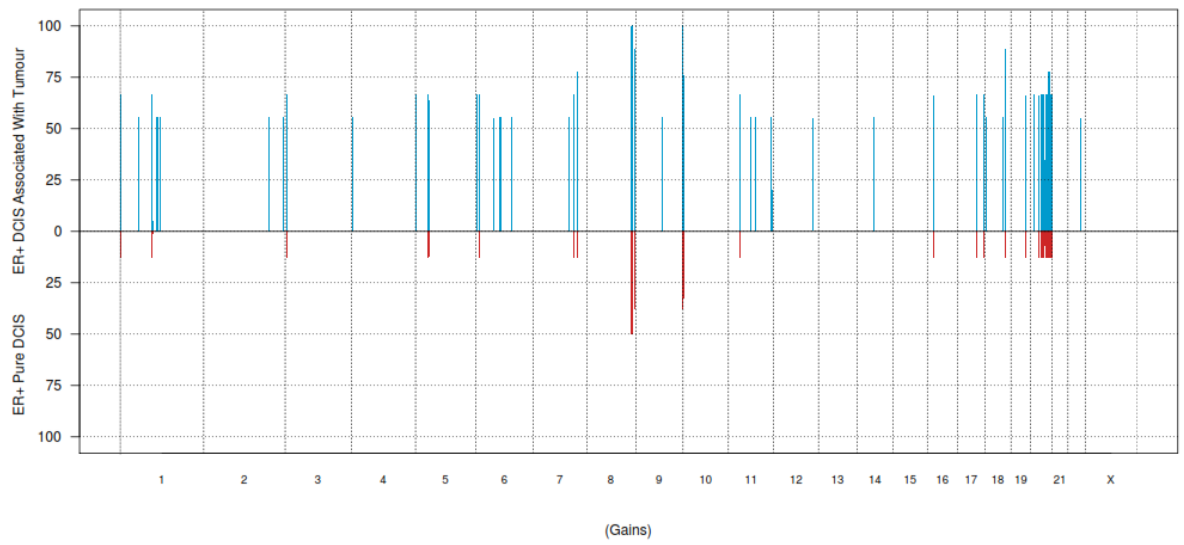
Frequency Plots_ ER+ DCIS Associated With Tumour vs ER+ Pure DCIS



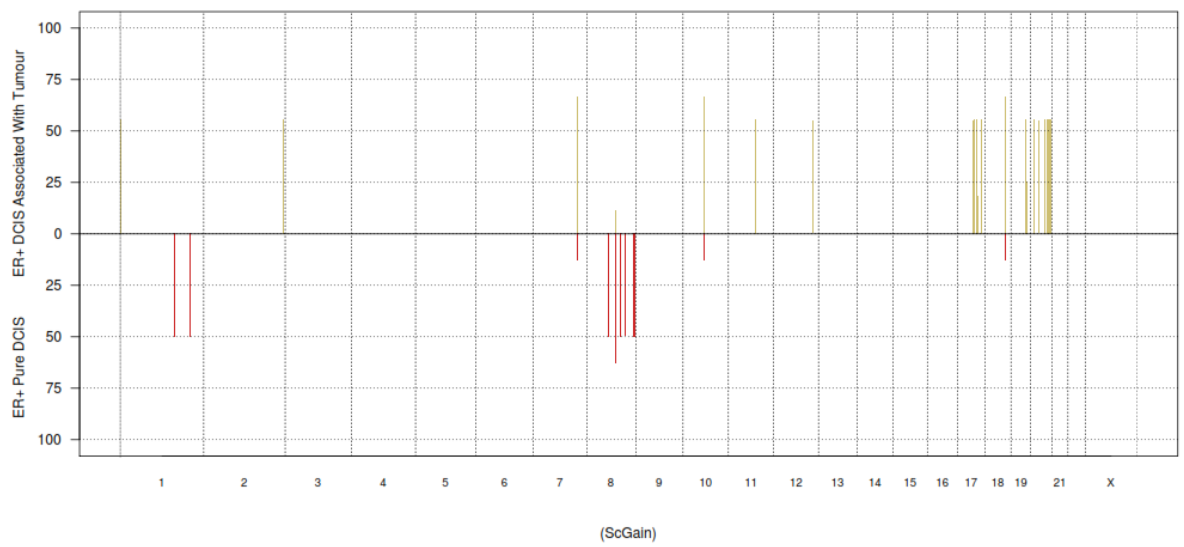
Frequency Plots_ ER+ DCIS Associated With Tumour vs ER+ Pure DCIS

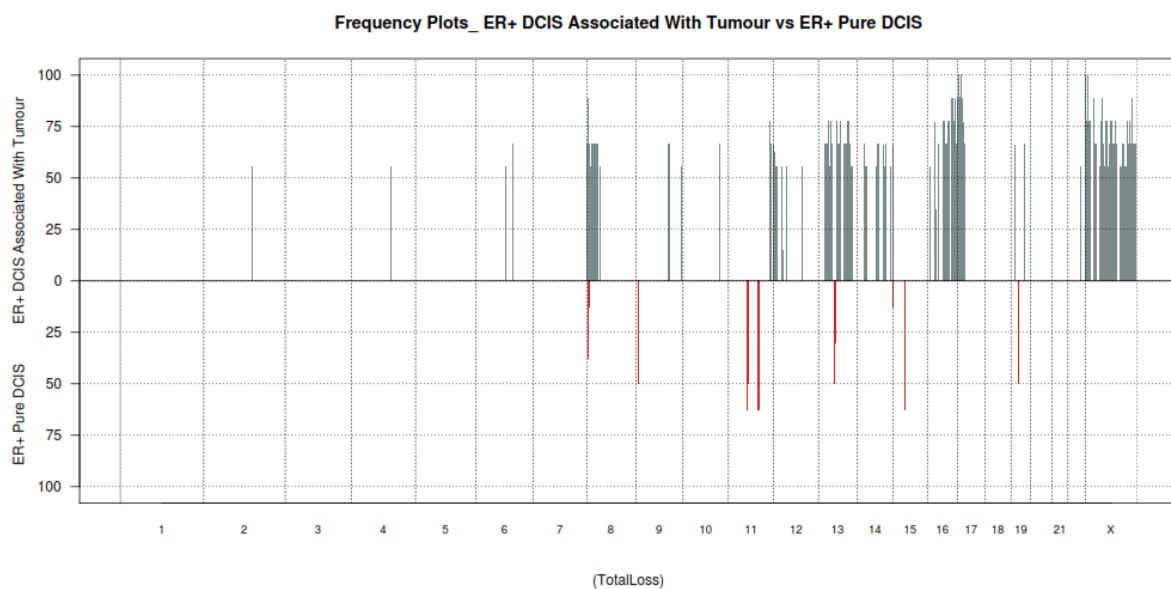
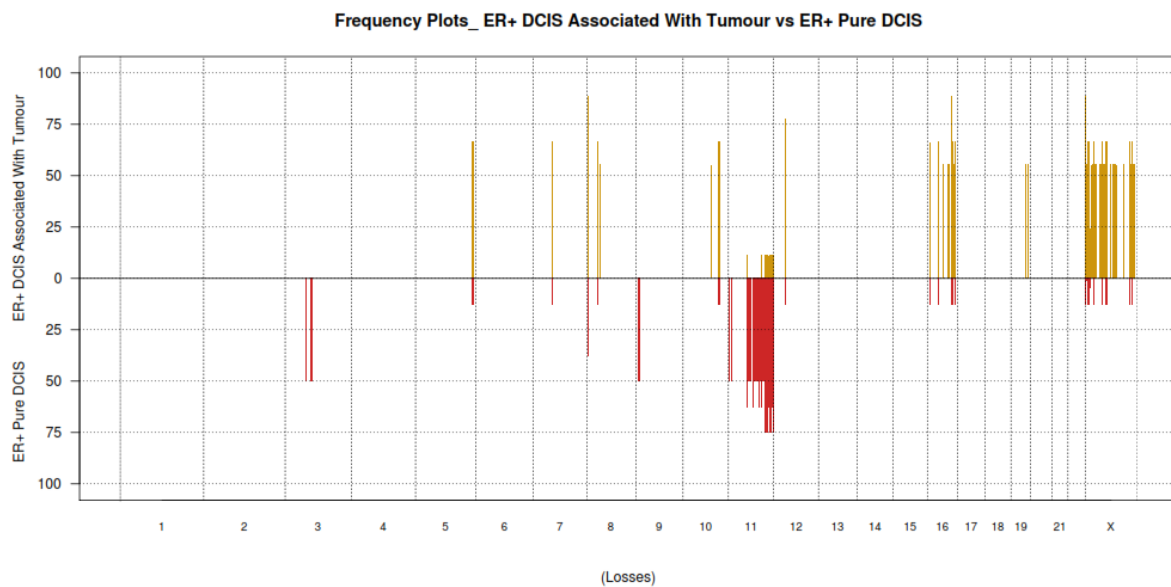


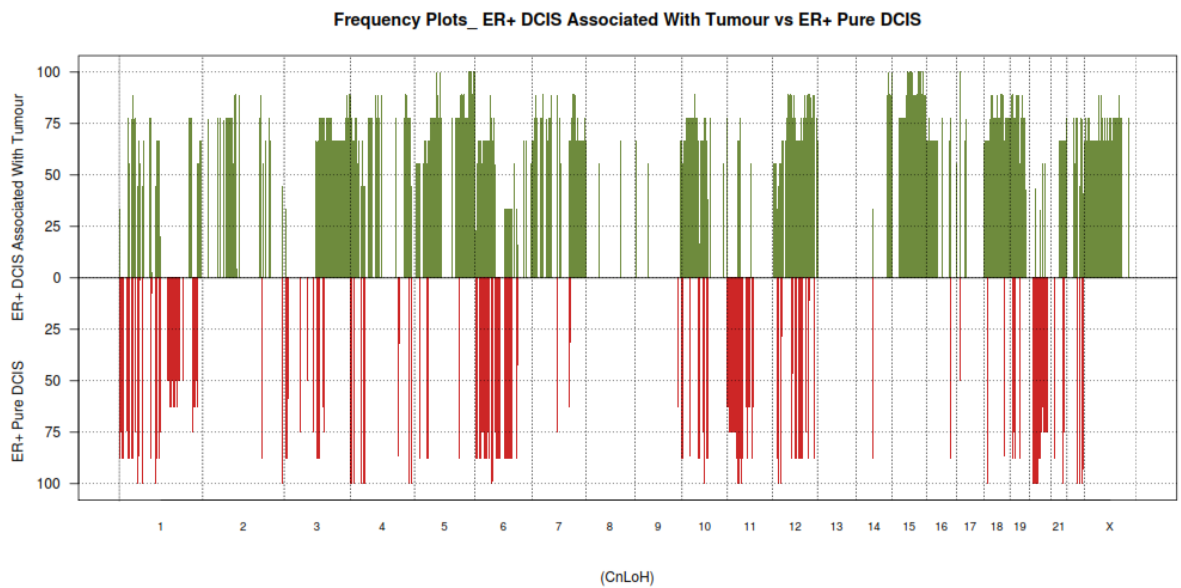
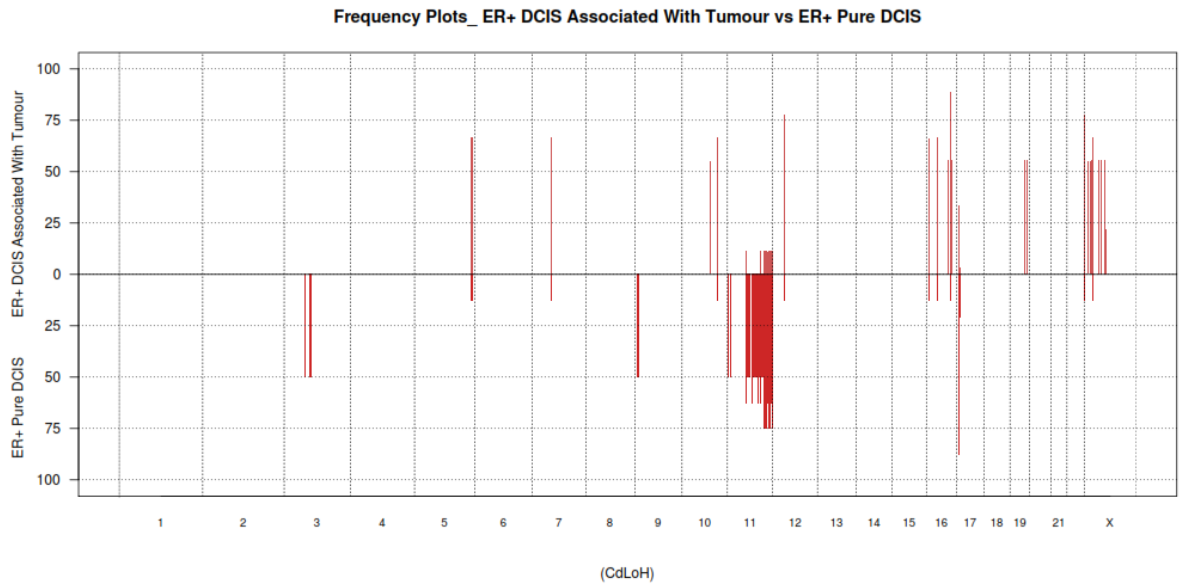
Frequency Plots_ ER+ DCIS Associated With Tumour vs ER+ Pure DCIS



Frequency Plots_ ER+ DCIS Associated With Tumour vs ER+ Pure DCIS







5.8.4.1 Amplification in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease

There were no amplifications in ER positive pure DCIS (n=0/8) in this series. However, there were amplifications present in cases of ER positive DCIS associated with invasive breast disease (n=5/9).

1. p-values < 0.05;
2. A gene list is mapped from 7 genomic regions found on chromosomes 6, 8.

- These regions encompass 69 genes altered in ER positive DCIS associated with invasive breast disease;

Genes:

5S_rRNA AC025647.3 AC103863.1 AF130342.1 AL031777.1 ANKRD46 AP001207.1 AP001208.1 AP001208.2 AP002852.1
 AP002907.1 CHRAC1 DENND3 EIF2C2 FBXO43 GRHL2 HFE HIST1H1C HIST1H1D HIST1H1E HIST1H1PS1
 HIST1H1T HIST1H2AB HIST1H2AC HIST1H2AD HIST1H2AE HIST1H2APS3 HIST1H2APS4 HIST1H2BB
 HIST1H2BC HIST1H2BD HIST1H2BE HIST1H2BF HIST1H2BG HIST1H2BH HIST1H2BI HIST1H3A HIST1H3B HIST1H3C
 HIST1H3D HIST1H3E HIST1H3F HIST1H3G HIST1H4A HIST1H4B HIST1H4C HIST1H4D HIST1H4E HIST1H4F
 HIST1H4G hsa-mir-1273 hsa-mir-151 KB-1615E4.1 NCALD PTK2 RGS22 RNF19A RP1-221C16.7 RPS10P1 RRM2B
 SNORD77 snoU13 SPAG1 TRAPPC9 U6atac U7 U91328.2 UBR5 Y_RNA

PANTHER analysis: 29 mapped ids are found, 40 mapped ids are not found.

5.8.4.2 Duplication in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease

There were no duplications present in ER positive DCIS associated with invasive breast disease in this series. However, there were duplications identified in ER positive pure DCIS (n=4/8).

- p-values < 0.05;
- A gene list is mapped from 1 genomic region and chromosome 8;
- These regions encompass 2 genes altered in ER positive pure DCIS;

Gene: ASPH 7SK

PANTHER analysis: 1 mapped id is found.

Gene ID: ENSG00000198363

Protein ID: [Q12797](#)

Gene Name: Aspartyl/asparaginyl beta-hydroxylase

Gene Symbol(s): ASPH

5.8.4.3 Genomic Gains in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease

There were gains associated with both ER positive pure DCIS (n=4/8) and ER positive DCIS associated with invasive breast disease (n=9/9).

1. p-values < 0.05;
2. A gene list is mapped from 43 genomic regions found on chromosomes 1, 3, 5, 6, 7, 10, 11, 16, 17, 18, 19, 20;
3. These regions encompass 283 genes altered in both ER positive pure DCIS and ER positive DCIS associated with invasive breast disease;

Genes:

5S_rRNA	7SK	AC005220.3	AC006076.1	AC006486.1	AC007431.1	AC007431.2	AC009362.2	AC013434.1	AC016831.7	
		AC018642.1	AC018816.3	AC058791.1	AC058791.2	AC090625.1	AC090821.2	AC090821.3	AC090821.4	AC090821.5
										AC090955.3
		AC103706.1	ACAP3	ADIG	AGRN	AL021394.1	AL021578.1	AL096828.1	AL109823.2	AL109840.1
										AL118506.1
		AL118506.3	AL121673.1	AL121673.2	AL121827.1	AL133330.1	AL136532.1	AL139348.1	AL162505.1	AL365229.1
										AL390719.1
		AL669831.3	AL713922.2	ARF4P2	ARFGAP1	ARFRP1	ARL6IP1	ATP5E	AURKAIP1	B3GALT6
										BCAS1
		BCAS4	BCL2	BCL2L1	BHLHE23	BIRC7	C1orf159	C1orf170	C20orf11	C20orf135
										C20orf166
		C20orf181	C20orf195	C20orf20	C20orf200	C20orf83	C20orf90	CANT1	CBFA2T2	CD44
										CDH4
		CHRNA4	CIC	COL20A1	COL9A3	COX4I2	CPSF3L	CST3	CTD-3184A7.4	CTSZ
										DBNDD2
		DENND3	DIDO1	DIP2C	DNAJC5	DPH3B	DVL1	EDN3	EEF1A2	EGOT
										ERF
										F3
		FAM132A	FAM41C	FAM65C	GATA5	GLTPD1	GMEB2	GNAS	GNASAS	GPR20
										GSK3A
		HAR1A	HAR1B	HAR1F	HES4	HM13	hsa-mir-1-1	hsa-mir-124-3	hsa-mir-133a-2	hsa-mir-1914
		hsa-mir-200a	hsa-mir-200b	hsa-mir-296	hsa-mir-298	hsa-mir-429	hsa-mir-647	hsa-mir-941-1	hsa-mir-941-3	hsa-
		mir-941-4	ID1	IL9RP2	ISG15	ITPR1	KCND2	KCNQ2	KDSR	KIAA0406
										KLHL17
										LGALS3BP
										LIME1
		MKLN1	MRPS16P	MSI2	MXRA8	NCOA3	NCRNA00029	NKAIN4	NOC2L	NTSR1
		NXT1	OGFR	OSMR	PAFAH1B3	PARD6B	PDHX	PHACTR3	PIGT	PLEKHN1
										PPDPF
		PRPF6	PTK6	PUSL1	RALGAPB	RBPJL	RICTOR	RP11-164D18.2	RP11-206L10.13	
		RP11-243J16.7		RP11-261N11.8		RP11-305P22.5		RP11-358D14.2	RP11-429E11.2	
		RP11-429E11.3		RP11-465B22.3		RP11-465B22.5		RP11-465B22.6	RP11-54O7.1	
		RP11-54O7.10		RP11-54O7.11		RP11-54O7.14		RP11-54O7.2	RP11-54O7.3	
		RP11-631M21.1		RP11-631M21.2		RP11-631M21.6	RP11-86H7.6	RP11-93B14.4	RP11-93B14.5	
		RP11-93B14.6		RP1-1J6.2	RP1-290I10.2	RP1-290I10.3		RP1-290I10.6	RP1-290I10.7	RP1-
		290I10.8		RP1-309F20.2	RP1-309F20.3	RP13-30A9.1		RP13-30A9.2	RP3-322G13.7	RP3-
		324O17.2	RP3-324O17.4		RP3-324O17.5	RP3-453C12.8		RP3-461P17.10	RP4-530I15.6	RP4-
		543J19.8		RP4-564F22.2		RP4-564F22.5		RP4-583P15.10	RP4-583P15.11	RP4-614C15.2
		RP4-614C15.3		RP4-697K14.12		RP4-697K14.3		RP4-697K14.7	RP4-719C8.1	RP4-
		724E16.2		RP4-797C5.2	RP4-806M20.2	RP5-1049G16.4		RP5-1100H13.3	RP5-827L5.1	RP5-827L5.2
										RP5-
		885L7.10		RP5-885L7.11		RP5-890O3.3		RP5-890O3.9	RP5-902P8.10	RP5-963E22.4
		RPL7P3	RPRD1B	RPS3P2	RTSL1	SAMD10	SAMD11	SCNN1D	SDC4	SDF4
										SFRS1
		SLC17A9	SLC1A2	SLC2A4RG	SLC45A4	SLCO4A1	SLMO2	SMG1	SNHG11	SNORA40
										SNORA71
		SNORA71A	SNORA71B	snoU13	SNTA1	SRMS	ST3GAL1	STMN3	SULF2	SUMO1P1
										SYCP2
		SYS1	TAF4	TAS1R3	TCFL5	TFAP2A	TGM2	TH1L	TMSL6	TNFRSF18
										TNFRSF4

TNFRSF6B	TOP1	TP53TG5	TPD52L2	TPX2	TSPAN12	TLL10	TUBB1	U1	U6	U7
UBE2J2	UCKL1	VSTM2L	WFDC2	Y_RNA	YTHDF1	ZBTB46	ZGPAT	ZMYND11		ZNF217
ZNF512B	ZNF831									

PANTHER analysis: 131 mapped ids are found, 152 mapped ids are not found.

There were gains present in ER positive DCIS associated with invasive breast disease (n=5/9) not observed in ER positive pure DCIS (n=0/8).

1. p-values < 0.05
2. A gene list is mapped from 35 genomic regions found on chromosomes 1, 2, 4, 5, 6, 7, 9, 11, 12, 14, 18, 19, 20, 22;
3. These regions encompass 147 genes altered in ER positive DCIS associated with invasive breast disease;

Genes:

5S_rRNA	7SK	AC000032.1	AC010442.1	AC012305.1	AC022724.2	AC026740.1	AC027097.1	AC064834.2	AC064834.3
AC064834.4	AC091609.1	AGAP1	AHRR	AIM1	AL031963.1	AL031963.2	AL031963.3	AL031963.5	AL109754.1
AL117692.1	AL117692.2	AL133351.1	AL136164.1	AL139092.1	AL161626.1	AL356735.1	AMIGO1	AMPD2	ANKRD5
AP000769.1	AP000944.1	AP002478.2	AP002478.4	ATG5	ATP5F1	ATP8B1	ATXN7L2	BPHL	C14orf182
C14orf183	C5orf55	C6orf155	C6orf195	CCDC127	CELSR2	CEP72	CTD-2083E4.4		CTD-2083E4.5
CTD-2083E4.6		CTD-2228K2.1		CTD-2228K2.2		CTD-2228K2.5		CTD-2231H16.1	
CTD-2589H19.4		CYB561D1	DLGAP1	ETS1	EXOC3	FAM136B	FRMD8	GCNT1	GNAI3
GNAT2	GPR61	GSK3A	GSTM2	GSTM3	GSTM4	GSTM5	hsa-mir-197	hsa-mir-30a	hsa-mir-30c-2
hsa-mir-612	KDSR	LRRC14B	MYBPHL	MYLK4	NQO2	OGFRL1	OVGP1	PDCD6	PLEKHG4B
PRDM1	PRUNE2	PSMA5	PSRC1	RIPK1	RP11-101C11.1		RP11-109I2.3		RP11-145H9.3
RP11-154D6.1		RP11-214N16.2		RP11-218C14.5		RP11-310P5.1		RP11-310P5.2	
RP11-404H14.1		RP11-416N4.1		RP11-416N4.4		RP11-420G6.3		RP11-420G6.4	
RP11-422N19.3		RP11-552M11.2		RP11-552M11.4		RP11-58E21.1		RP11-66C24.1	
RP11-811I15.1		RP11-94F13.1		RP1-40E16.11		RP1-40E16.2		RP1-40E16.3	RP1-
40E16.4	RP1-40E16.8		RP1-40E16.9		RP1-90J20.10		RP1-90J20.2	RP1-90J20.7	RP1-90J20.8
331H24.4	RP4-630J13.1		RP4-735C1.4		RP5-1160K1.1		RP5-1160K1.3		RP5-1160K1.6
	RP5-839B4.7		RP5-839B4.8		RSL24D1P11	SDHA	SERPINB1	SERPINB6	SERPINB9
	SLC39A10	SLC9A3	SNAP25	SORT1	SYPL2	TPPP	TUBB2A	U6	USP24
80423F1.1	WRNIP1	Y_RNA	ZDHHC11	ZDHHC11B	ZFYVE28				WDR77
									WI2-

5.8.4.4 Genomic Sc Gains in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease

There were Sc gains present in ER positive pure DCIS (n=4/8) not observed in any ER positive DCIS associated with invasive breast disease (n=0/9) in this series.

1. p-values < 0.05;
2. A gene list is mapped from 6 genomic regions found on chromosomes 1, 8;
3. These regions encompass 16 genes altered in ER positive pure DCIS;

Genes:

AC018442.1 AC023933.1 AL590138.1 CD244 CSMD3 F11R hsa-mir-151 ITLN1 ITLN2 PLXNA2 PTK2
RP11-312J18.5 RP11-312J18.6 RP11-312J18.7 RP11-544M22.1 VPS13B

PANTHER analysis: 8 mapped ids are found, 8 mapped ids are not found.

There were Sc gains present in ER positive DCIS associated with invasive breast disease (n=6/9) not observed in ER positive pure DCIS (n=0/8).

1. p-values < 0.05;
2. A gene list is mapped from 23 genomic regions found on chromosomes 1, 2, 11, 12, 17, 19, 20;
3. These regions encompass 143 genes altered in ER positive DCIS associated with invasive breast disease;

Genes:

ABCC3 AC005220.3 AC007431.1 AC012305.1 AC091062.2 AC092066.6 ACAP3 AGAP1 AGRN AL049736.1
AL109754.1 AL109840.1 AL136532.1 AL139348.1 AL365229.1 AL390719.1 AL669831.3 ANKRD5 ATP5E AURKAIP1
B3GALT6 BCAS1 BCAS4 BCL3 C1orf73 C1orf159 C1orf170 C20orf177 C20orf83 CBLC
CDH26 CDH4 CPSF3L CTB-22K21.2 CTSZ CYP24A1 DOK5 DVL1 FAM132A
FAM41C FAM65C GLTPD1 GNAS HES4 HOXB3 HOXB4 HOXB5 hsa-mir-10a hsa-mir-200ahsa-
mir-200b hsa-mir-429 ISG15 KLHL17 LUC7L3 MRPS16P MSI2 MXRA8 NDC2L NXT1 PARD6B
PFDN4 PHACTR3 PLEKHN1 PPP1R3D PTPRT PUSL1 RNFT1 RP11-164D18.2 RP11-206L10.13
RP11-294J22.5 RP11-357H14.7 RP11-416N4.1 RP11-416N4.4 RP11-429E11.2
RP11-429E11.3 RP11-465B22.3 RP11-465B22.5 RP11-465B22.6 RP11-506D12.1
RP11-506D12.3 RP11-54O7.1 RP11-54O7.10 RP11-54O7.11 RP11-54O7.14 RP11-54O7.2
RP11-54O7.3 RP11-700H6.1 RP11-700H6.2 RP1-232N11.2 RP1-309F20.2 RP1-309F20.3
RP3-322G13.7 RP4-530I15.6 RP4-543J19.8 RP4-583K8.1 RP4-614C15.2 RP4-614C15.3
RP4-715N11.2 RP4-719C8.1 RP4-723E3.1 RP4-724E16.2 RP4-730D4.1 RP5-1022J11.1
RP5-1022J11.2 RP5-1030M6.2 RP5-1073F15.1 RP5-827L5.1 RP5-827L5.2 RP5-839B4.7 RP5-839B4.8

RP5-890O3.3	RP5-890O3.9	RP5-902P8.10	RPL12P4 RPS6KB1	SAMD11	SCNN1D
SDF4	SLC39A11	SLMO2	SNAP25 SNORA70	snoZ6	SYCP2
TNFRSF18	TNFRSF4 TOB1	TTLL10	TUBB1	U4atac	U6
Y_RNA ZFP64	ZNF217	ZNF831	U7	UBE2J2	WFIKKN2

PANTHER analysis: 62 mapped ids are found, 79 mapped ids are not found.

5.8.4.5 Losses in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease

There is some overlap in the losses found in ER positive pure DCIS (n=5/8) compared to ER positive DCIS associated with invasive breast disease (n=1/9). DCIS.

1. p-values < 0.05;
2. A gene list is mapped from 35 genomic regions found on chromosomes 3, 9, 11;
3. These regions encompass 510 genes altered in ER positive pure DCIS;

Genes:

5S_rRNA	7SK	ABCG4	AC015689.1	AC019227.1	AC019227.2	AC025029.1	AC087441.1	AC090587.2	AC090587.4
			AC090587.5	AC090804.1	AC108448.2	AC108448.3	AC109309.4	AC132217.4	ACA59
			ACAD8	ADAMTS15	ADAMTS8				
AMICA1	AMOTL1	ANKK1	ANKRD42	ANKRD49	AP000560.1	AP000619.1	AP000620.1	AP000620.2	AP000642.1
			AP000654.1	AP000654.2	AP000757.1	AP000767.1	AP000769.1	AP000787.1	AP000797.1
			AP000797.2	AP000843.2	AP000844.1				
			AP000866.1	AP000866.2	AP000873.1	AP000880.1	AP000892.4	AP000907.1	AP000907.2
			AP000907.3	AP000908.1	AP000911.1				
AP000925.2			AP000936.1	AP000936.3	AP000936.4	AP000936.5	AP000944.1	AP001007.1	AP001007.3
			AP001007.3	AP001122.1	AP001267.1				
			AP001267.2	AP001528.1	AP001582.1	AP001767.1	AP001781.1	AP001830.1	AP001891.1
			AP001922.1	AP001972.1	AP001979.1				
			AP001992.2	AP001999.1	AP002364.1	AP002364.2	AP002380.1	AP002380.2	AP002451.1
			AP002451.1	AP002754.1	AP002783.1	AP002783.2			
AP002783.3			AP002783.4	AP002783.5	AP002783.6	AP002802.1	AP002884.4	AP002886.2	AP002956.4
			AP002956.4	AP002962.1	AP002986.1				
			AP002991.1	AP003025.2	AP003039.3	AP003039.4	AP003041.1	AP003070.1	AP003072.1
			AP003122.1	AP003175.1	AP003304.2				
			AP003392.1	AP003392.3	AP003402.1	AP003402.2	AP003461.1	AP003461.10	AP003461.11
			AP003461.2						
			AP003461.3	AP003461.4	AP003461.5	AP003461.6	AP003461.7	AP003461.8	AP003461.9
			AP003499.3	AP003500.1	AP003774.4				
			AP003774.5	AP003774.7	AP003780.2	AP003781.1	AP003971.1	AP004242.1	AP004242.2
			AP004248.4	AP004372.1	AP004607.1				
AP004607.2			AP004607.3	AP004609.2	AP004609.4	AP005273.1	AP005435.1	AP005597.1	AP005638.1
			AP005638.3	AP005718.1					
			AP005814.1	AP006216.10	AP006216.11	AP006216.12	AP006216.5	AP006288.1	APLP2
			APOA1						
APOA4	APOA5	APOC3	ARCNI	ARHGAP32	ARHGAP42	ARHGEF12	ARHGEF17	ARRB1	ATM
ATP5L	B3GAT1	BACE1	BARX2	BCL9L	BCO2	BEST1	BIRC2	BIRC3	BTG4
			BUD13						
C11orf1	C11orf10	C11orf2	C11orf30	C11orf36		C11orf45	C11orf52	C11orf53	C11orf54
			C11orf57						
C11orf61	C11orf63	C11orf65	C11orf66	C11orf71	C11orf75	C11orf82	C11orf84	C11orf88	C11orf9
C11orf90	C11orf92	C11orf93	CADM1	CAPN1	CARS	CBL	CCDC15	CCDC153	CCDC67
CCDC82	CCDC83	CCDC84	CCDC89	CCDC90B	CD3D	CD3E	CD3G	CDC42EP2	CEP164
			CEP57						
CHEK1	CHORDC1	CHRM1	CLDN25	CNTN5	CREBZF	CRTAM	CRYAB	CTSC	CUL5
			CWC15						
CXCR5	DAGLA	DCUN1D5	DDI1	DDX25	DDX6	DGAT2	DIXDC1	DLAT	DLG2
DPF2	DRD2	DSCAML1	DYNC2H1	EI24	EIF3F	ENDOD1	ESAM	ETS1	EXPH5
FADS1	FADS2	FADS3	FAM168A	FAM181B	FAM55A	FAM55B	FAM55D	FAM76B	FAU
FCHSD2	FDX1	FDXACB1	FEN1	FLI1	FOLH1B	FOLR4	FOXR1	FRMD8	FTH1
			FUT4						

	FXYD2	FXYD6	FZD4	GBE1	GLB1L2	GLB1L3	GPR83	GUCY1A2	HEPACAM	HEPN1	
	HRASLS5	hsa-mir-1304	hsa-mir-1908	hsa-mir-34b	hsa-mir-34c	hsa-mir-483	hsa-mir-548l	hsa-mir-611	hsa-mir-612	HSPB2	HTR3A
HTR3B	HYLS1	IFT46	IGF2	IGF2AS	IGSF9B	IL10RA	IL18	INS	JAM3	JRKL	
	KCNJ1	KCNJ5	KDELC2	KDM4D	KDM4DL	KIAA1731	LAYN	LRRRC10B	MALAT1	MAML2	
	MAP6	MARK2	ME3	MED17	MLL	MMP1	MMP10	MMP12	MMP13	MMP20	MMP27
	MMP3	MMP7	MMP8	MOGAT2	MPZL2	MPZL3	MRE11A	MRGPRE	MRGPRG		MRPL49
	MTMR2	MYLKP1	NAALAD2	NCAM1	NCAPD3	NCRNA00167	NEU3	NFRKB	NLRX1	NNMT	
	NOX4	NPAT	NRGN	NUP98	OPCML	OR10V1	OR2AT2P	OR2AT4	OR4A15	OR4A16	
	OR55B1P	OR8A1	OR8B12	OR8B9P	OSBP	OSBPL5	P2RY2	P2RY6	PAFAH1B2	PANX3	PATL1
	PCF11	PCSK7	PDGFD	PDZD3	PGAP2	PGR	PHLDB1	PICALM	PIH1D2	PIWIL4	
	PKNOX2	POLA2	POU2AF1	PPIHP1	PPP2R1B	PRCP	PRDM10	PTPRD	PUS3	RAB30	
	RAB38	RAB39	RAB3IL1	RBM7	RDX	RELT	REXO2	RHOG	RNF169	RNF214	
	ROBO3	ROBO4	RP11-142L1.1	RP11-142L1.2		RP11-142L1.3		RP11-147I3.1		RP11-160H12.1	
	RP11-237N19.1		RP11-358N4.2		RP11-359D24.1		RP11-399J13.1		RP11-438N5.1		
	RP11-447G14.1		RP11-567M21.1		RP11-657K20.1		RP11-659G9.1		RP11-667M19.1		
	RP11-690D19.1		RP11-727A23.1		RP11-742N3.1		RP11-810P12.1		RP11-832N8.1		
	RP11-861M13.1		RP11-864G5.1		RP11-959F10.1		RPL23AP64	RPL37AP8	RPLP0P2	RRM1	
SCARNA11	SCN2B	SCN4B	SDHD	SESN3	SFRS2B	SIAE	SIDT2	SIK2	SIK3	SLC22A10	
	SLC22A20	SLC22A24	SLC22A25	SLC22A9	SLC25A45	SLC35F2	SLC36A4	SLC37A2	SLCO2B1	SNORA1	
SNORA18	SNORA25	SNORA32	SNORA40	SNORA7	SNORA70	SNORA8	SNORD112	SNORD5	SNORD56		
	snoU13	snoU2_19	snoU2-30	snoZ40	SORL1	SPA17	SPATA19	SPCS2	SPDYC	ST14	STIM1
	STT3A	SYT7	SYTL2	SYVN1	TAF1D	TAGLN	TBRG1	TEX12	THYN1	TIGD3	
	TIMM8B	TM7SF2	TMEM123	TMEM126A	TMEM126B	TMEM135	TMEM136	TMEM218	TMEM25	TMEM45B	
	TMPRSS13	TMPRSS4	TMPRSS5	TP53AIP1	TREH	TRIM48	TRIM49	TRIM49L	TRIM53	TRIM53B	
	TRIM64	TRIM64B	TRIM77	TTC12	TTC36	U2	U6	U7	UBASH3B	UBE4A	UBTFL2
	UBTFL3	UPK2	USP28	UVRAG	VPS26B	VSIG2	XRRA1	Y_RNA	ZBTB16	ZBTB44	
ZC3H12C	ZNF259	ZNHIT2	ZW10								

PANTHER analysis: 286 mapped ids are found, 224 mapped ids are not found.

There were losses present in ER positive DCIS associated with invasive breast disease (n=6/9) not observed in ER positive pure DCIS (n=0/8).

1. p-values < 0.05;
2. A gene list is mapped from 56 genomic regions found on chromosomes 5, 8, 10, 16, 19, X;
3. These regions encompass 149 genes altered in ER positive DCIS associated with invasive breast disease;

Genes: 5S_rRNA AC002366.3 AC003666.1 AC009055.1 AC009110.1 AC009120.1 AC009161.1 AC010546.1 AC018558.1 AC022164.1
AC023824.2 AC025287.1 AC025287.2 AC025287.3 AC025287.4 AC069278.1 AC073493.1 AC073493.2 AC073493.3 AC135776.1
ADAM3A ADAM5P AGTR2 AKR1B1P8 AL035443.1 AL133512.1 AL353698.1 AL591398.1 AMELX ANKRD26P1
ARHGAP6 ARHGFE9 ASPDH BMP15 BX119964.1 BX119964.2 BX119964.3 BX119964.4 C5orf50 CCNYL3 CDH11
CDH8 CLCN4 CTD-2503H21.1 CXorf29 DHRSX DMD EFHC2 EIF2S3 FAAH2 FAM48B2 FDP5L5

	FMR1	FMR1NB	FRMPD4	GPR112	GPR64	GS1-256O22.1	GS1-256O22.5	GS1-466O4.2			
	GS1-466O4.3		GS1-466O4.5	GS1-541M1.2		JOSD2	KLF8	L29074.1	L29074.2	L29074.3	
	L29074.5	LRRC4B	MAGEB18	MAGEB5	MAGEB6	MAGEB6B	MAGED4	MAGEE1	MAP3K7IP3	NRG3	
	PSMD7	RP11-104D21.1		RP11-104D21.2	RP11-104D21.3		RP11-121C22.1		RP11-156J23.1		
	RP11-1A15.2		RP11-357C3.2		RP11-357C3.3		RP11-363G10.2		RP1-137H15.2	RP1-	
138A5.1	RP11-38O23.3		RP11-38O23.4		RP11-38O23.5		RP11-38O23.7		RP11-3D23.1		
	RP11-434J24.2		RP11-467P22.4		RP11-479I1.4	RP11-517D11.3		RP11-54I5.1	RP11-552E4.2		
	RP11-617O8.1		RP11-637B23.1		RP11-637B23.3		RP11-702C7.1		RP11-702C7.2		
	RP11-761E20.1		RP11-81C12.1		RP11-93B10.3		RP1-20J23.1	RP1-267M5.1	RP1-267M5.2	RP1-	
290C9.2	RP13-188A5.1		RP13-346H10.2		RP13-348B13.1	RP13-348B13.2		RP13-34C21.1		RP13-60P5.2	
	RP1-54B20.7		RP3-433G13.1	RP4-551E13.2		RP5-1145L23.2		RP5-849L7.1	RP6-1O2.1	SHROOM4	
SLITRK4	SNORA11	SNORA68	snoU13	SPACA5B	SPANXN3	SPIN4	SSX5	SSX6	U40455.1	U6	VENTXP1
	WVOX	Y_RNA	Z98304.1	ZFP112	ZFRP1	ZFX	ZIC3	ZNF180	ZNF229	ZNF285	

PANTHER analysis: 42 mapped ids are found, 106 mapped ids are not found.

5.8.4.6 Total Loss in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease

There were total losses present in ER positive pure DCIS (n=5/8) not observed in ER positive DCIS associated with invasive breast disease (n=0/9).

1. p-values < 0.05;
2. A gene list is mapped from 13 genomic regions found on chromosomes 9, 11, 13, 15, 19;
3. These regions encompass 25 genes altered in for ER positive pure DCIS;

Genes:

7SK	AC025678.1	AC027139.2	AC027139.3	AP000720.1	AP001482.1	AP003072.1	CCDC67	CTSC	GOLGA8B	GRM5	hsa-
mir-1233	KDM4C	LPAR6	NOX4	OR10V1	OSBP	PATL1	RB1	RP11-657K20.1		SLC36A4	TYR
	U7	ZNF208	ZNF257								

PANTHER analysis: 15 mapped ids are found, 10 mapped ids are not found.

There were total losses present in ER positive DCIS associated with invasive breast disease (n= 9/9) not observed in ER positive pure DCIS (n=0/8).

1. p-values < 0.05;
2. A gene list is mapped from 91 genomic regions found on chromosomes 11, 13, 16, 17, X;

3. These regions encompass genes altered in ER positive DCIS associated with invasive breast disease;

Genes:

5S_rRNA	7SK	ABCC12	ABR	AC000003.1	AC002347.1	AC002347.2	AC002366.3	AC003658.1	AC004070.1										
				AC004478.2	AC005277.1	AC005284.1	AC005323.1	AC005358.1	AC005358.3	AC005548.1	AC005548.2	AC005696.2	AC005696.3						
				AC005725.1	AC006435.1	AC007225.1	AC007339.1	AC009053.1	AC009110.1	AC009120.1	AC009120.3	AC009161.1	AC010546.1						
				AC015799.1	AC015799.2	AC015908.2	AC016292.1	AC016292.3	AC025287.1	AC025287.2	AC025287.3	AC025287.4	AC025518.1						
				AC027045.1	AC027763.1	AC027763.2	AC032044.1	AC055839.2	AC073493.1	AC073493.2	AC073493.3	AC087498.1	AC087501.1						
				AC087742.1	AC087742.2	AC090282.1	AC090617.1	AC091153.4	AC097370.1	AC099508.1	AC099684.1	AC100748.2	AC104581.1						
AC109599.1	AC116914.1		AC118754.2	AC118754.4	AC127521.1	AC130343.2	AC130689.5	AC135178.7	ACA64	ADAT1									
AIPL1	AL136160.1	AL355852.1	AL390738.2	AL450226.1	AL450226.2	AL589987.1	ALOX12	ALOX15	AMAC1L3										
AMELX	ANKFY1	ANKRD26P1		AP001970.1	ARHGAP6	ARHGEF15	ARHGEF9	ARRB2	ASB12	ASGR1	ASGR2								
ASMT	ASPA	ATP2A3	ATP5EP2	BCAR1	BCL6B	BHLHA9	BMP15	BRWD3	BX004827.1	BX649443.1									
C13orf36	C16orf78	C16orf87	C17orf100	C17orf44	C17orf48	C17orf49	C17orf68	C17orf85	C17orf91										
CAMKK1	CBLN1	CCDC42	CDH13	CDH8	CFDP1	CHD3	CHRNB1	CHST5	CHST6										
CLEC10A	CLEC18B	CRK	CRLF2	CTD-2503H21.1		CTD-2545G14.1		CTNS	CTRB1	-CTRB2									
CYB5D1	CYB5D2	DDX26B	DHRS7C	DHRSX	DLG4	DNAH2	DNAH9	DNAJA2	DPH1	EEF1A1L1									
EIF4A1P10	ERN2	FA2H	FAM123B	FAM46D	FAM64A	FANCB	FBXO39	FRMPD4	GABARAPL2	GAPDHP34									
GAS7	GGT6	GLG1	GLP2R	GLRA2	GPT2	GS1-466O4.2	GS1-466O4.3	GS1-466O4.5											
GS1-541M1.1		GS1-541M1.2		GSG2	GTPBP6	HCCS	HDHD1A	HIC1	HK2P1										
HMGA1L1	hsa-mir-1253	hsa-mir-132	hsa-mir-1468	hsa-mir-195	hsa-mir-212	hsa-mir-22	hsa-mir-497	INPP5K	ITGAE	KARS									
KDM6B	KIAA0664	KIAA0753	KRBA2	LDHD	LL0YNC03-3F3.1	LSMD1	MAGEB18	MAGEB5	MAGEB6										
MAGEB6B	MAGED1	MAGED4	MAGED4B	MAGEE1	MAP2K4	MED31	METT10D	MFSD6L	MLKL	MNT									
MTMR8	MTUS2	MYBBP1A	MYH1	MYH10	MYH13	MYH2	MYH3	MYH4	MYH8										
MYLK3	MYO1C	MYOCD	NCRNA00086	NCRNA00107		NCRNA00108		NDEL1	NETO2										
NPIPL2	NTN1	NXN	ODF4	OR1A1	OR1A2	OR1D2	OR1D5	OR1E1	OR1E2	OR1G1									
OR3A1	OR3A2	OR3A3	OR3A4	ORC6L	OVCA2	P2RX1	P2RX5	P2RY8	PAFAH1B1	PAN3									
PCDH11X	PDX1	PELP1	PFAS	PHKB	PIK3R5	PIK3R6	PIRT	PITPNA	PITPNM3										
PLCXD1	PLK1	POLR2A	PPP2R3B	PRPF8	PRPSAP2	PSMD7	RAB43P1	RANGRF	RAP1GAP2										
RCVRN	RFWD3	RILP	RNASEK	RNF222	RP1-107N3.1	RP1-107N3.2		RP11-114H20.1											
RP11-120D5.1		RP11-145B3.2	RP11-151A2.2		RP11-16L6.3	RP11-197L7.2		RP11-217H19.1											
RP11-234P3.2	RP11-234P3.3		RP11-234P3.4		RP11-234P3.5		RP11-265D19.5		RP11-265D19.6										
RP11-284B18.3		RP11-2N21.2		RP11-309M23.1		RP11-309M23.2	RP11-311P8.1		RP11-311P8.2										
RP11-314A20.1		RP11-314A20.2	RP11-325E14.1		RP11-325E14.2		RP11-325E14.5		RP11-331F4.1										
RP11-331F4.4	RP11-33N11.1		RP11-342C20.3		RP11-344N17.11		RP11-344N17.2		RP11-344N17.3										
RP11-344N17.6		RP11-344N17.7		RP11-344N17.8		RP11-357C3.2		RP11-357C3.3											
RP11-360A9.1		RP11-363G10.2		RP11-364B14.1		RP11-364B14.2		RP11-364B14.3											
RP11-403E24.1		RP11-432N13.2		RP11-432N13.3		RP11-459A10.1		RP11-459A10.2											
RP11-459A10.3		RP11-462C21.1		RP11-471M2.1		RP11-471M2.2		RP11-481F23.1											
RP11-483M24.2		RP11-48B14.1		RP11-51C14.1		RP11-51C14.2		RP11-520F9.2											
RP11-552E4.3		RP11-552E4.4		RP11-552E4.5	RP11-565F19.1		RP11-572F4.1		RP11-609D21.1										
RP11-617O8.1		RP11-637B23.1		RP11-637B23.3		RP11-647I17.1		RP11-702C7.1											
RP11-702C7.2		RP11-713H12.1		RP11-93B10.3		RP11-963H4.1		RP11-963H4.2		RP1-									
279N11.1	RP13-210D15.1		RP13-210D15.4		RP13-210D15.6		RP13-297E16.3		RP13-297E16.4										
	RP13-297E16.5		RP13-34C21.1		RP13-465B17.2		RP13-465B17.3		RP13-565O16.2										
	RP13-60P5.2		RP13-858C7.1		RP13-926M18.1		RP3-428A13.1		RP4-710L4.1		RPA1								
RPL26	RTN4RL1	SCARF1	SCARNA21	SCO1	SERPINF1	SERPINF2	SFRS17A	SGSM2	SHCBP1										

SHISA6	SHOX	SHPK	SHROOM4	SLC13A5	SLC16A11	SLC16A13	SLC25A35	SLC43A2	SLITRK5	SMG6
SMTNL2	SMYD4	SNORA11	SNORA69	SNORA76	SNORD91	snoU13	SOX3	SPATA22	SPDYE4	
SPIN4	SPNS2	SPNS3	SRR	SSX1	SSX3	SSX9	STX8	TAX1BP3	TBX22	TEKT1
TERF2IP	TIMM22	TLCD2	TMEM170A		TMEM220	TMEM231	TMEM88	TMEM93	TNFSF12	TNFSF12-TNFSF13
TNFSF13	TRPV1	TRPV3	TSR1	TUSC5	TXNDC17	U1 U6	U7	U70984.1	U8	
UBE2G1	UPRT	USP43	VAT1L	VENTXP1	VPS35	WDR16	WDR59	WDR81	WSCD1	XAF1
Y_RNA	YWHAE	ZBTB4		ZDHHC15	ZFP1	ZNF423	ZNF449	ZNF75D	ZNRF1	ZZEF1

PANTHER analysis: 219 mapped ids are found, 223 mapped ids are not found.

5.8.4.7 CdLOH in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease

There is CdLOH present in ER positive pure DCIS (n=6/8) not observed in any ER positive DCIS associated with invasive breast disease (n=0/9).

1. p-values < 0.05;
2. A gene list is mapped from 54 genomic regions found on chromosomes 3, 9, 11;
3. These regions encompass 333 genes altered in ER positive pure DCIS;

Genes:

5S_rRNA	ABCG4	AC015689.1	AC025029.1	AC087441.1	AC090587.2	AC090587.4	AC090587.5	AC090804.1	AC108448.2	AC108448.3	
	AC109309.4	AC132217.4	AC133041.1	ACAD8	AMOTL1	ANKRD42	ANKRD49	AP000438.1	AP000560.1	AP000619.1	
	AP000620.1	AP000620.2	AP000642.1	AP000654.1	AP000654.2	AP000767.1	AP000769.1	AP000787.1	AP000873.1	AP000911.1	
	AP000944.1	AP001267.1	AP001267.2	AP001528.1	AP001767.1	AP001922.1	AP001972.1	AP001979.1	AP001992.2	AP001999.1	
	AP002364.1	AP002364.2	AP002380.1	AP002380.2	AP002754.1	AP002783.1	AP002783.2	AP002783.3	AP002783.4	AP002783.5	
	AP002783.6	AP002802.1	AP002956.4	AP002991.1	AP003122.1	AP003175.1	AP003304.2	AP003392.1	AP003392.3	AP003461.1	
	AP003461.10	AP003461.11	AP003461.4	AP003461.5	AP003461.7	AP003499.3	AP003774.4	AP003774.5			
	AP003774.7	AP003780.2	AP003781.1	AP003971.1	AP004242.1	AP004242.2	AP004607.1	AP004607.2	AP004607.3	AP004609.2	
	AP004609.4	AP005273.1	AP005435.1	AP005718.1	AP005814.1	AP006288.1	ARCNI	ARHGAP42	ARHGEF12	ARHGEF17	
	ARRB1	ATM	ATP5L	B3GAT1	BCL9L	BEST1	BIRC2	BIRC3	C11orf10	C11orf2	C11orf30
C11orf36	C11orf54	C11orf65	C11orf66	C11orf75	C11orf82	C11orf84	C11orf9	C11orf90	CADM1	CAPN1	
	CARS	CBL	CCDC153	CCDC67	CCDC82	CCDC83	CCDC84	CCDC89	CCDC90B	CD3G	
	CDC42EP2	CEP57	CHORDC1	CHRM1	CNTN5	CREBZF	CTD-2026G6.1	CTD-2026G6.2		CWC15	
	CXCR5	DAGLA	DCUN1D5	DDI1	DDX6	DGAT2	DLG2	DPF2	DYNC2H1	EIF3F	
	ENDOD1	EXPH5	FADS1	FADS2	FADS3	FAM168A	FAM181B	FAM76B	FAM86D		FAU
	FCHSD2	FEN1	FOLH1B	FOLR4	FOXRI	FRMD8	FTH1	FUT4	FZD4	GBE1	GLB1L2
	GLB1L3	GPR83	GUCY1A2	HRASLS5	hsa-mir-1304	hsa-mir-1324	hsa-mir-1908	hsa-mir-483	hsa-mir-548l	hsa-mir-611	hsa-
mir-612	IFT46	IGF2	IGF2AS	IGSF9B	INS	JAM3	JRKL	KDELC2	KDM4D	KDM4DL	
	KIAA1731	LRRC10B	MALAT1	MAML2	MAP6	MARK2	ME3	MED17	MLL	MMP1	MMP10
	MMP12	MMP13	MMP20	MMP27	MMP3	MMP7	MMP8	MOGAT2	MRE11A	MRGPRE	
MRGPRG	MRPL49	MTMR2	MYLKP1	NAALAD2	NCAPD3	NEU3	NLRX1	NOX4	NPAT	NUP98	OR2AT2P
	OR2AT4	OR55B1P	OSBPL5	P2RY2	P2RY6	PCF11	PDGFD	PDZD3	PGAP2	PGR	PHLDB1
	PICALM	PIWIL4	POLA2	PRCP	PTPRD	RAB30	RAB38	RAB3IL1	RELT	RHOG	RNF169
	RP11-142L1.1	RP11-142L1.2	RP11-142L1.3	RP11-147I3.1	RP11-160H12.1	RP11-241K7.1					
	RP11-241K7.2	RP11-358N4.2	RP11-359D24.1	RP11-399J13.1	RP11-413E6.1	RP11-413E6.2					

RP11-413E6.3	RP11-438N5.1	RP11-447G14.1	RP11-642N14.3	RP11-659G9.1	
RP11-690D19.1	RP11-727A23.1	RP11-742N3.1	RP11-803B1.1	RP11-803B1.2	
RP11-803B1.3	RP11-803B1.4	RP11-803B1.5	RP11-803B1.7	RP11-810P12.1	
RP11-861M13.1	RP11-864G5.1	RP11-959F10.1	RPL23AP64 RPLP0P2	RRM1	SES3
SFRS2B	SLC22A10 SLC22A20	SLC22A24	SLC22A25	SLC22A6	SLC22A8
					SLC22A9 SLC25A45
SNORA1	SNORA18	SNORA25	SNORA32	SNORA40	SNORA7
					SNORA70
					SNORA8
SNORD56	snoU13	snoU2_19	snoU2-30	snoZ40	SPATA19 SPCS2 SPDYC
					STIM1
					SYT7
TAF1D	THYN1	TIGD3	TM7SF2	TMEM123	TMEM126A
					TMEM126B
TRIM49	TRIM49L	TRIM53	TRIM53B	TRIM64	TRIM77
				TTC36	U2
					U6
					UBE4A
UBTF3	UPK2	UVRAG	VPS26B	XRRA1	Y_RNA
					ZNHIT2

PANTHER analysis: 167 mapped ids are found, 166 mapped ids are not found.

There is CdLOH found in ER positive DCIS associated with invasive breast disease.

(n=8/9) with some overlap in ER positive pure DCIS (n=1/8).

1. p-values < 0.05;
2. A gene list is mapped from 36 genomic regions found on chromosomes 5, 7, 10, 12, 16, 19, X;
3. These regions encompass 61 genes, 2 and 59 genes altered in ER positive DCIS associated with invasive breast disease;

Genes:

AC002519.2	AC005774.2	AC009055.1	AC009120.1	AC018558.1	AC022164.1	AC023824.2	AC025287.2	AC069278.1	AC073057.5
AC135776.1	ARHGEF9	ASPDH	BX649443.1	C16orf67	C5orf50	CCNYL3	CDH11	DHRX	EFHC2
									EIF2S3
FAM48B2	FDP5L5	FRMPD4	GPR64	JOSD2	LRRC4B	MAGED4	NRG3	RP11-114G11.2	RP11-1217F2.15
RP11-133M24.1	RP11-357C3.2	RP11-357C3.3	RP11-363G10.2	RP11-479I1.4					
RP11-617O8.1	RP11-93B10.3	RP1-290C9.2	RP13-297E16.4	RP13-297E16.5	RP13-34C21.1				
RP13-858C7.1	RP4-551E13.2	SNORA11	SNORA68	SORCS3	SPIN4	U7	VN1R66P	Y_RNA	
ZFP112	ZFX	ZNF180	ZNF229	ZNF285	ZNF716	ZNF720	ZNF720P1		

PANTHER analysis: 41 mapped ids are found, 28 mapped ids are not found.

5.8.4.8 CnLOH in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease

There is CnLOH present in ER positive pure DCIS (n=8/8) not observed in ER positive DCIS associated with invasive breast disease (n=0/9).

1. p-values < 0.05;

2. A gene list is mapped from 98 genomic regions found on chromosomes 1, 2, 3, 4, 6, 10, 11, 12, 16, 19, 20, 21;
3. These regions encompass 627 genes altered in ER positive pure DCIS;

Genes:

5S_rRNA	7SK	ABI3BP	ABT1	AC000032.1	AC002485.1	AC004699.1	AC012154.1	AC018809.6	AC021749.1	
	AC021749.2	AC021749.3	AC022007.1	AC022007.2	AC022007.4	AC022007.5	AC023085.1	AC025423.2	AC027288.1	AC027288.2
	AC034102.1	AC034193.5	AC034193.7	AC063962.1	AC073366.1	AC078917.1	AC079598.1	AC079598.2	AC087521.1	AC087521.2
	AC087521.3	AC087521.4	AC087521.5	AC093028.1	AC096649.1	AC096649.2	AC096649.3	AC124890.1	AC131263.1	AC131263.2
	AC137628.1	AC137834.2	ACCS	ACCSL ADAM30	ADAMTS14	ADH5P4	AF254982.1	AF254982.2	AF254982.3	
AGAP10	AGAP6	AGAP7	AGAP8	AHCYL1	AL021807.2	AL021808.1	AL021808.2	AL021808.3	AL023913.1	
	AL031178.1	AL035045.1	AL035456.1	AL035467.1	AL035661.3	AL050321.1	AL050321.2	AL050322.1	AL050335.1	AL050335.2
	AL109754.1	AL109922.1	AL118509.1	AL133255.1	AL136164.1	AL138850.1	AL139429.1	AL354680.1	AL354933.1	AL355344.1
AL359752.1	AL391137.2	AL442003.2	AL442003.3	AL450342.1	AL450342.2	AL512288.1	AL512503.1	AL590062.1	AL591034.2	
	AL591044.1	AL591044.2	AL591044.3	AL591044.4	AL672187.3	AL672187.4	AL713998.1	ALKBH3 ALX3 AMBRA1	ANKRD5	
ANKRD52	ANO3	ANXA8L1	AP000302.58	AP000303.1	AP000304.12		AP000304.2	AP001992.2	APOF	
ARHGAP1	ARHGAP9	ASAH2	ATF2	ATP5G3	B3GAT2	BANF2	BDNF	BDNFOS	BMS1P2	BTBD3
BYSL	C10orf53	C11orf91	C1orf127	C1orf167	C1orf216	C20orf12	C20orf187	C20orf3	C20orf61	
C20orf72	C20orf78	C20orf79	C20orf94	C22orf30	C3orf10	C3orf24	C6orf155	C6orf205	C6orf57	
CASP14	CCDC34	CCT2	CD59	CD82	CDK2 CFLP6		CHRM4	CHST1	CIDECP	
CLCN6	CLSPN	CNPY2	COMTD1	COQ10A	CR381653.1	CR381653.2	CR381653.3	CR381670.1	CR381670.2	
CR382285.2	CR382287.1	CR392039.1	CR392039.2	CR392039.3	CREB3L1	CRELD1	CRYZL1	CS	CSF1	
CSRP2BP	CST1	CST2	CST3	CST4	CST5	CST7	CSTF3	CSTP1	CSTP2	CTA-215D11.3
CTA-215D11.4	CTSLL7	CYCSP41	DEFB110	DEFB112	DEFB113	DEFB114	DEFB133	DEK	DEPDC7	
DGKZ	DIP2C	DNAJC28	DONSON	DPCR1	DTNBP1	DUPD1	DUSP13	ELOVL5	ENO1	
EPHA6	EPS8L3	ERBB3	ERRFI1	ESYT1	EXT2	EYA3	EYS	F2	FAM135A	
FAM21A	FAM21D	FAM25B	FAM25D	FAM35B2	FAM40A	FAM72B	FANCD2	FAT1P1	FBXO3	
FBXO9	FCGR1B	FRS2	FRS3	GART	GCLC	GCM1	GGTLA4	GLI1	GLUDP8	GMPR
GPR182	GSTM2	GSTM3	GSTM3P	GSTM5	GUSBL1	HARBI1	HCG21	HIPK3	HIST1H2AH	
HIST1H2BJ	HIST1H2BK	HIST1H4I	HIST2H2BA	HIST2H2BB	HMGAL1L5	HNRNPM	hsa-mir-1228	hsa-mir-1252	hsa-mir-129-2	
hsa-mir-30a	hsa-mir-30c-2	hsa-mir-670	hsa-mir-933	HSD17B12	ICK	IFNGR2	IGHV1OR15-5	IKZF4	IL17RC	
IL23A	INHBC	INHBE	IRAK2	ITSN1	JAG1	JARID2	KDM1B	KHDRBS2	KIAA0319L	
KIAA0652	KIAA1274	KIF16B	LGR4	LIN28AP3	LIN7C	LRP1	LRRC10	MACROD2	Mar02	MDK
MDM2	MED20	MKKS	MRPS11P1	MTHFR	MYCL1	MYL6	MYL6B	MYLIP	MYO1A	MYST4
NAB2	NCDN	NCOA4	NCRNA00240	NDUFA4L2	NEU3	NKX22	NKX2-4	NODAL	NOTCH2	NTS
NXPH4	NXT1	OBFC2B	OGDHL	OGFRL1	OTOR	OVOL2	PA2G4	PA2G4P2	PAK7	PAN2
PANK4	PARG	PARK7	PGAM3P	PGC	PHACTR1	PHF21A	PIP5K2A	PISD	PLCH2	
PLK1S1	POM121L1P	POM121L2	POM121L3P	PPIAP13	PPIAP17	PPYR1	PRAM1	PRF1	PRICKLE4	PRPF19
PRRT3	PRSS16	PSMB2	PTMAP3	QSER1	R3HDM2	RAB5B	RASSF9	RDH16	RERE	
RFX4	RIC8B	RIMS1	RNF144B	RNF41	RP11-102J14.1	RP11-104F15.7	RP11-109G10.2			
RP11-134K13.2		RP11-144G6.10		RP11-144G6.4		RP11-144G6.9	RP11-146I2.1			
RP11-154D6.1		RP11-157E14.1		RP11-157E14.3	RP11-159H22.1	RP11-160A9.1	RP11-160A9.2			
RP11-185B14.1	RP1-118J21.5		RP11-195M16.1		RP11-209A2.1	RP11-218C14.5	RP11-227D2.3			
RP11-239L20.3		RP11-239L20.4		RP11-239L20.5		RP11-244C20.1	RP11-263F15.1			
RP11-281P11.1		RP11-292F22.3		RP11-292F22.4		RP11-292F22.5	RP11-298J23.5			
RP11-298J23.6		RP1-129L7.1	RP11-307F22.2		RP11-314P12.2	RP11-314P12.3	RP11-324H6.5			
RP11-324H6.6		RP11-324H6.7		RP11-348A11.1		RP11-397G17.1	RP11-399K21.2			

RP11-399K21.4	RP11-399K21.5	RP11-399K21.6	RP1-13D10.2	RP1-13D10.3	RP1-13D10.4
RP1-13D10.4	RP1-13D10.5	RP11-401E14.1	RP11-401E14.2	RP11-416N4.1	RP11-416N4.4
RP11-425L10.1	RP11-431K24.1	RP11-431K24.2	RP11-431K24.3	RP11-431K24.4	RP11-435D7.3
RP11-439A17.10	RP11-439A17.2	RP11-439A17.4	RP11-439A17.5	RP11-439A17.7	
RP11-439A17.9	RP11-448N11.1	RP11-448N11.2	RP11-457M11.2	RP11-460I13.5	
RP11-460I13.6	RP11-462G2.1	RP11-462G2.2	RP11-466G11.1	RP11-472I20.1	
RP11-481A12.2	RP11-481A12.5	RP11-481A12.6	RP11-487I5.4	RP11-506E9.1	RP11-507K13.4
RP11-507K13.6	RP11-526K17.2	RP11-526K21.2	RP11-528A10.1	RP11-528A10.2	RP11-542F9.1
RP11-560J1.1	RP11-56A21.4	RP11-56N19.5	RP11-575L16.2	RP11-587D21.1	RP11-592B15.3
RP11-592B15.4	RP11-592B15.6	RP11-69I22.2	RP11-69I22.3	RP11-707M13.1	RP11-710A11.2
RP11-74E24.2	RP11-77G23.2	RP11-77G23.5	RP11-789A7.1	RP1-190J20.2	
RP11-96J19.1	RP11-96J19.2	RP11-96J19.3	RP1-198K11.3	RP1-20N4.1	RP1-240B8.3
RP1-278O22.1	RP1-278O22.2	RP1-27K12.4	RP1-288M22.1	RP1-288M22.2	RP1-298M8.1
RP13-15M17.1	RP13-392I16.1	RP1-34L19.1	RP1-34L19.2	RP1-89D4.1	RP3-322G13.7
RP3-395M20.2	RP3-395M20.3	RP3-410C9.1	RP3-410C9.2	RP3-477M7.2	RP3-477M7.3
RP3-477M7.4	RP3-483K16.2	RP3-483K16.3	RP3-527G5.1	RP4-564O4.1	RP4-568F9.3
RP4-568F9.6	RP4-568F9.7	RP4-580G13.1	RP4-633I8.1	RP4-633I8.2	RP4-633I8.3
RP4-633I8.4	RP4-697P8.2	RP4-717M23.2	RP4-718D20.3	RP4-726N1.2	RP4-728D4.2
RP4-728D4.3	RP4-734C18.1	RP4-735C1.4	RP4-735C1.5	RP4-735C1.6	RP4-737E23.2
RP4-737E23.4	RP4-742J24.2	RP4-777D9.2	RP4-796I8.1	RP5-1018A4.3	RP5-1042I8.6
RP5-1053E7.1	RP5-1053E7.2	RP5-1068E13.3	RP5-1068F16.3	RP5-1069C8.1	RP5-1069C8.2
RP5-1099D15.1	RP5-1100I6.1	RP5-1100I6.2	RP5-1115A15.1	RP5-1177M21.1	RP5-1195D24.1
RP5-839B4.7	RP5-839B4.8	RP5-858B16.5	RP5-860P4.2	RP5-872K7.2	RP5-872K7.7
RP5-905G11.3	RP5-931K24.1	RP5-973N23.4	RP5-983H21.3	RP5-984P4.1	RP5-984P4.4
RPL15P1	RPL23AP6	RPL24P2	RPL41	RPL7AL3	RPLP0P1
RPS11P1	RPS15AP1	RPS26	RPS5L	RRBP1	SAMD8
SDR9C7	SFTA2	SHMT2	SILV	SLC15A3	SLC24A3
SLC25A33	SLC39A5	SLC9A3P1	SLC9A3P3	SMAP1	SMARCC2
SMCHD1	SNAP25	SNORA48	SNORA70	SNORA74	SNORA77
SNORA8	SNORD17	snoU13	SNRBP2	SNX5	SON
SPCS2	SPDEF	SPTLC3	STAT2	STAT6	SUOX
SYT1	TAC3	TCP1	L1	TFAP2E	TFEB
TIMELESS	TIMM23	TIMM23B	TMEM109	TMEM111	TMEM132A
TMEM132B	TMEM194A	TMEM50B	TMEM90B	TOMM6	TP53I11
TPMT	TPTE	TRIT1	TSPAN18	TTC17	U1
U2	U3	U4	U6	U7	USP49
VARS	VDAC2	VN1R11P	VN1R12P	VN1R13P	VN1R14P
XRN2	XRRA1	XXbac-	BPG118E17.9	XXyac	YR14BB7.1
XXyac-YX60D10.1	Y_RNA	ZBTB39	ZC3H10	ZNF133	ZNF322A
ZNF408	ZPLD1				

PANTHER analysis: 248 mapped ids are found, 379 mapped ids are not found.

There is CnLOH in ER positive DCIS associated with invasive breast disease not observed in ER positive pure DCIS.

1. p-values < 0.05;
2. A gene list is mapped from 27 genomic regions found on chromosomes 5, 14, 15;

- These regions encompass 149 genes altered in ER positive DCIS associated with invasive breast disease;

Genes:

7SK AB019439.2 AC008659.1 AC008659.2 AC009997.1 AC010674.2 AC010724.1 AC010724.2 AC010724.3 AC010724.4 AC010724.5
AC010724.6 AC010724.7 AC011295.1 AC011295.2 AC011295.3 AC011295.4 AC011295.5 AC011944.1 AC012050.1 AC018926.1
AC018926.2 AC022748.1 AC025580.2 AC025917.1 AC026624.1 AC026956.1 AC026956.2 AC039057.1 AC039057.3 AC044907.1
AC068870.1 AC069082.1 AC084882.1 AC087738.2 AC092373.1 AC104115.1 AC105339.1 AC105339.2 AC116903.1 AC116903.2
AC126339.1 AC126339.2 AC126339.3 AC126339.4 AC135995.1 AC135995.2 AC135995.3 AC135995.4 AC135995.5 AC136940.3
ADAMT57 AGXT2L2 ANKRD34C AP3B2 ARPP19 BCL2L10 C15orf26 C15orf32 C15orf48 CCPG1
CHRNA4 CLK4 COL23A1 CPEB1 CTB-12904.1 CTD-2184D3.1 CTD-2270N23.1 CTSB CYP19A1
DYX1C1 EFTUD1 F2RL1 FAM153C FAM154B FAM174B FSD2 GFPT2 GNB5 GOLGA6L10
GOLGA6L3 GOLGA6L5 GOLGA6L9 HNRNPAB HNRNPB1 HOMER2 hsa-mir-1266 hsa-mir-147b hsa-mir-184 hsa-mir-628 IGHD
IL16 KIAA1024 KIAA1199 KIAA1370 MAPK9 MAST4 MESDC1 MESDC2 MEX3B MORF4L1 MYO5A
MYO5C NEDD4 NHP2 ODZ2 PIGB PRTG PYGO1 RAB27A RASGEF1C RASGRF1 RFX7 RMND5B
RP11-1001M11.1 RP11-252I14.1 RP11-287J9.1 RP11-335K5.1 RP11-5P22.1 RP11-5P22.2
RP11-5P22.3 RP11-889L3.1 RP11-889L3.3 RP11-889L3.4 RP11-96J23.1 RP13-608F4.1
RP13-996F3.1 RPS17 S100Z SCARNA15 SEMA6D SH3GL3 SLC30A4 snoU109 snoU13
SPATA5L1 ST8SIA2 STARD5 TEX9 TMC3 TMED3 U1 U2 U7 UBE2Q2P2 UNC13C
WDR72 WHAMM

PANTHER analysis: 64 mapped ids are found, 84 mapped ids are not found.

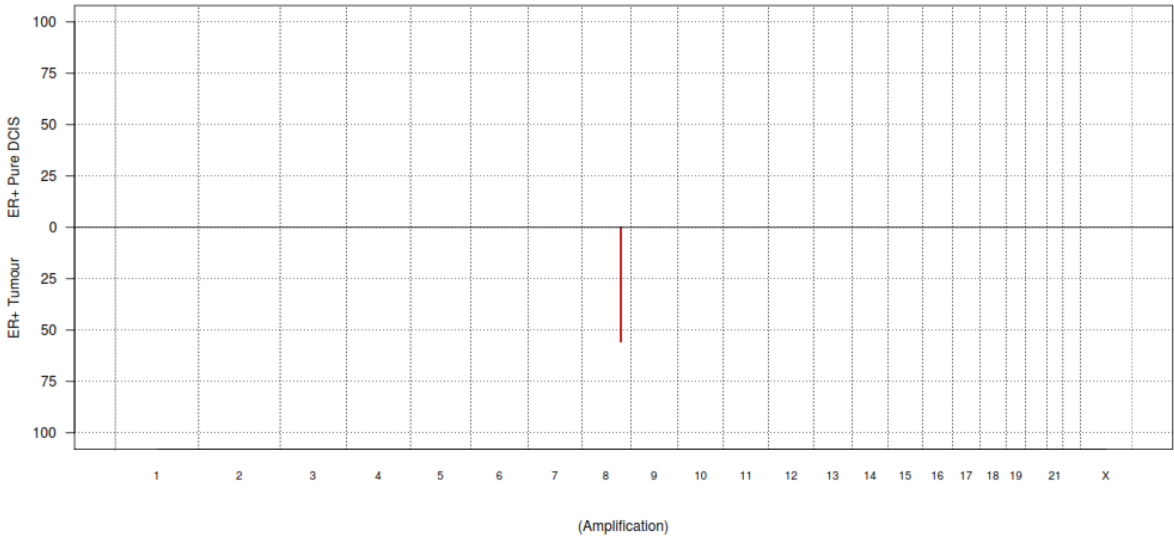
5.8.5 Copy Number Aberrations for Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Positive Invasive Breast Disease

These series examines the difference between oestrogen receptor positive pure DCIS and oestrogen receptor positive invasive breast cancer.

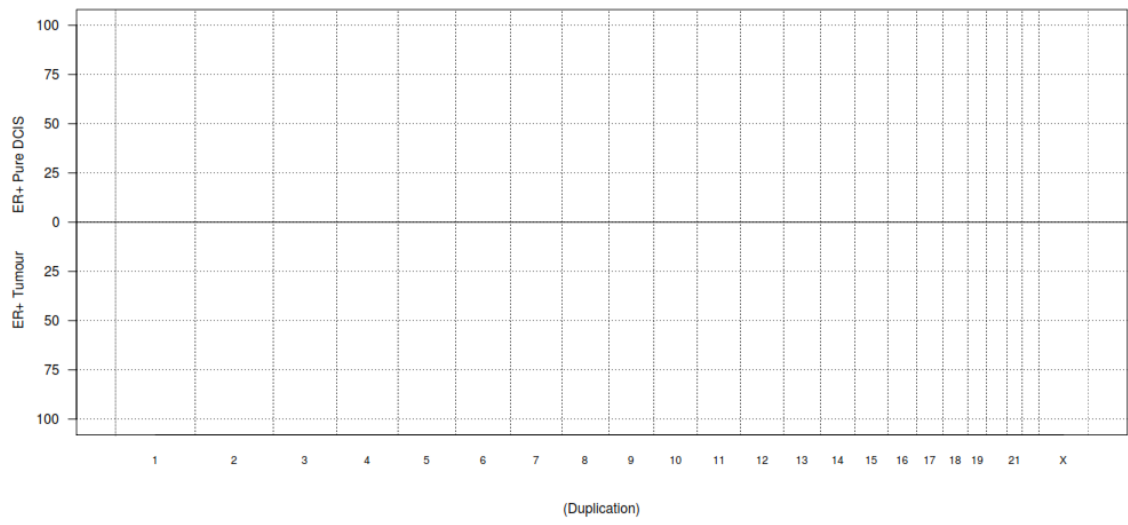
Frequency plots showing copy number aberrations between oestrogen receptor positive pure DCIS and oestrogen receptor positive invasive breast disease were provided by Breakthrough Breast Cancer/Research Oncology, King's College London Bioinformatics Department (Figure 38).

Figure 38: Frequency plots showing copy number aberrations between oestrogen receptor positive pure DCIS and oestrogen positive invasive breast disease (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 194-197).

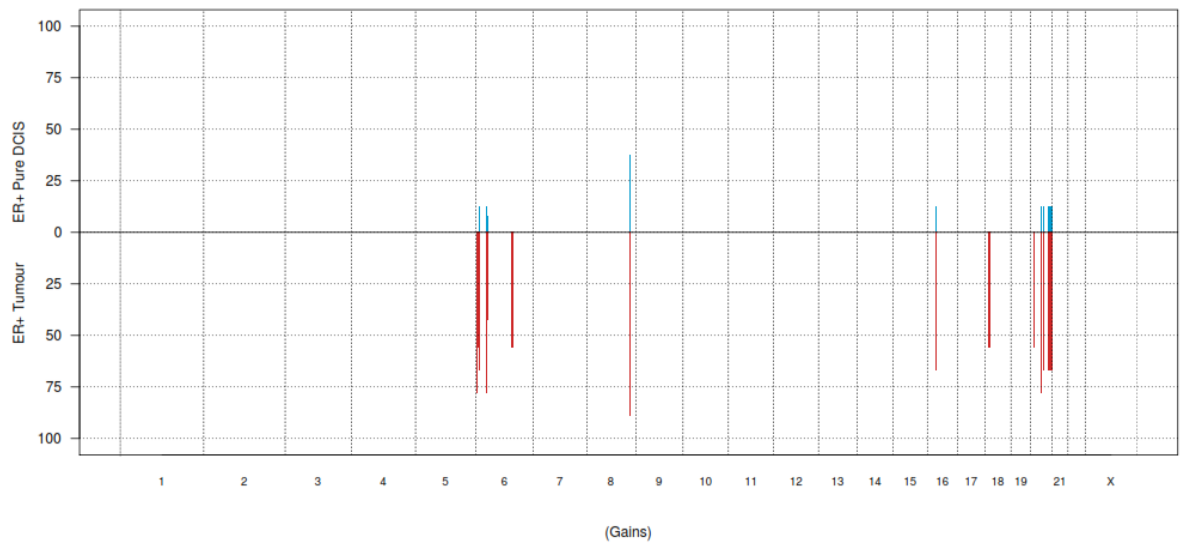
Frequency Plots_ ER+ Pure DCIS vs ER+ Tumour



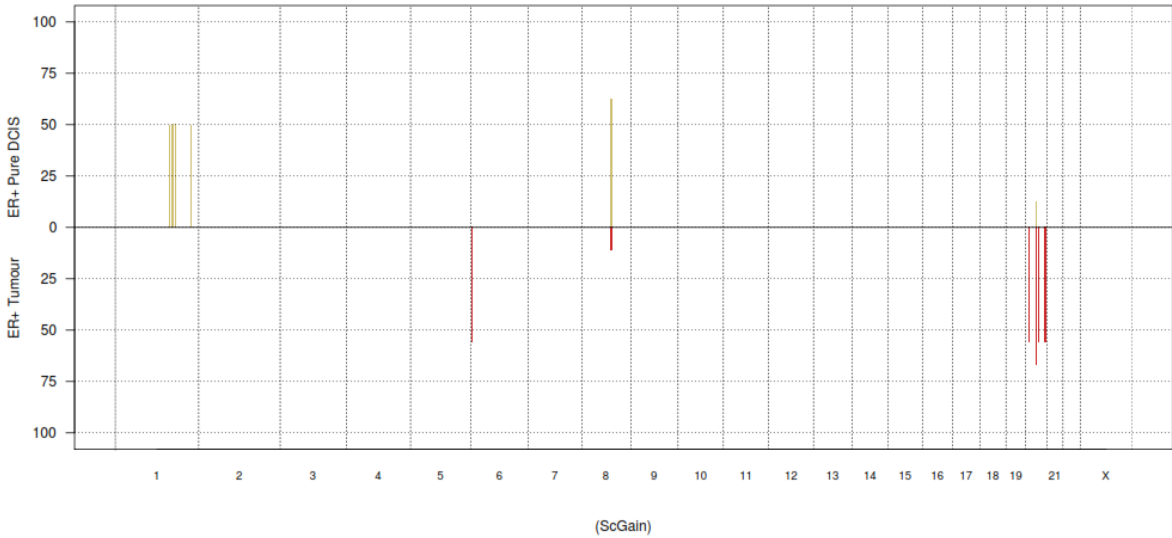
Frequency Plots_ ER+ Pure DCIS vs ER+ Tumour



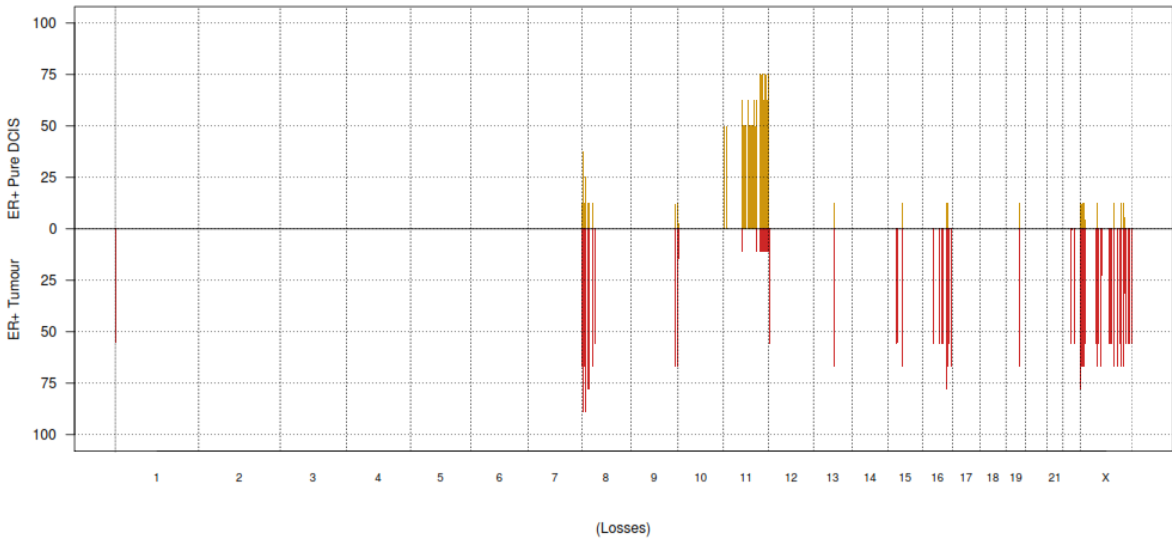
Frequency Plots_ ER+ Pure DCIS vs ER+ Tumour



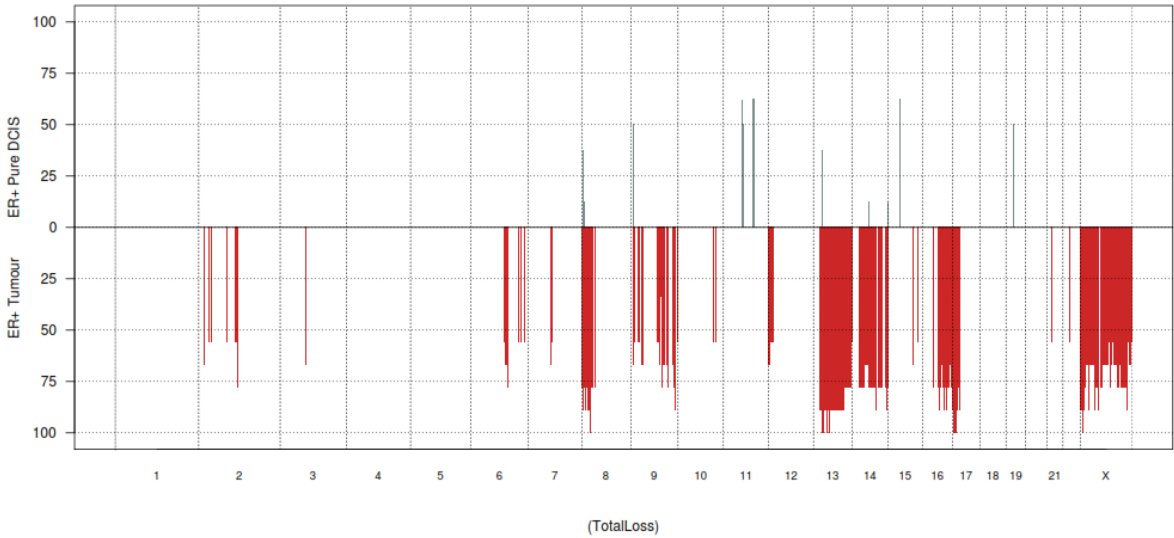
Frequency Plots_ ER+ Pure DCIS vs ER+ Tumour



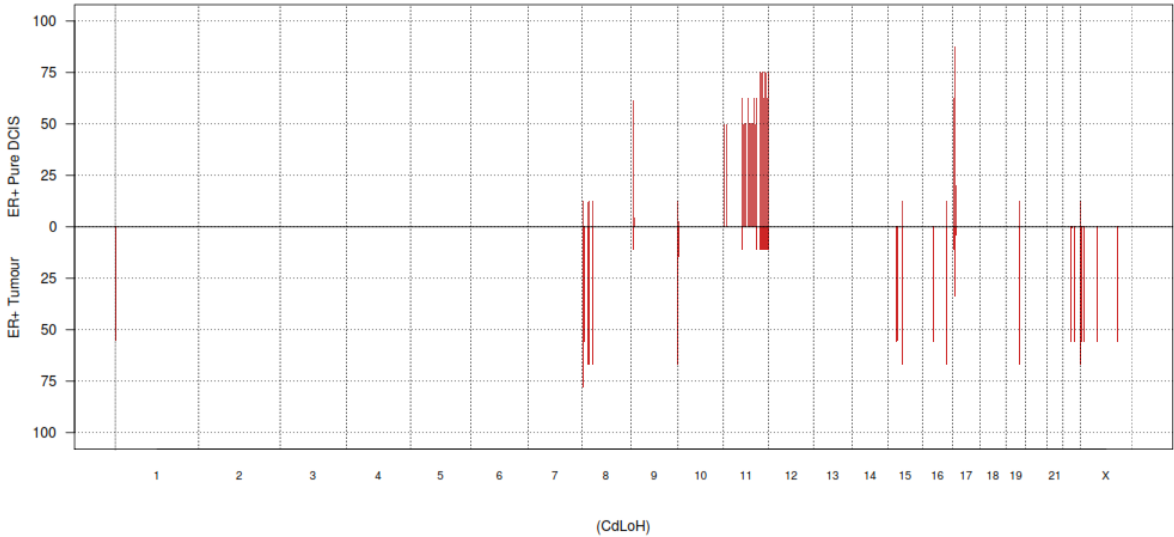
Frequency Plots_ ER+ Pure DCIS vs ER+ Tumour

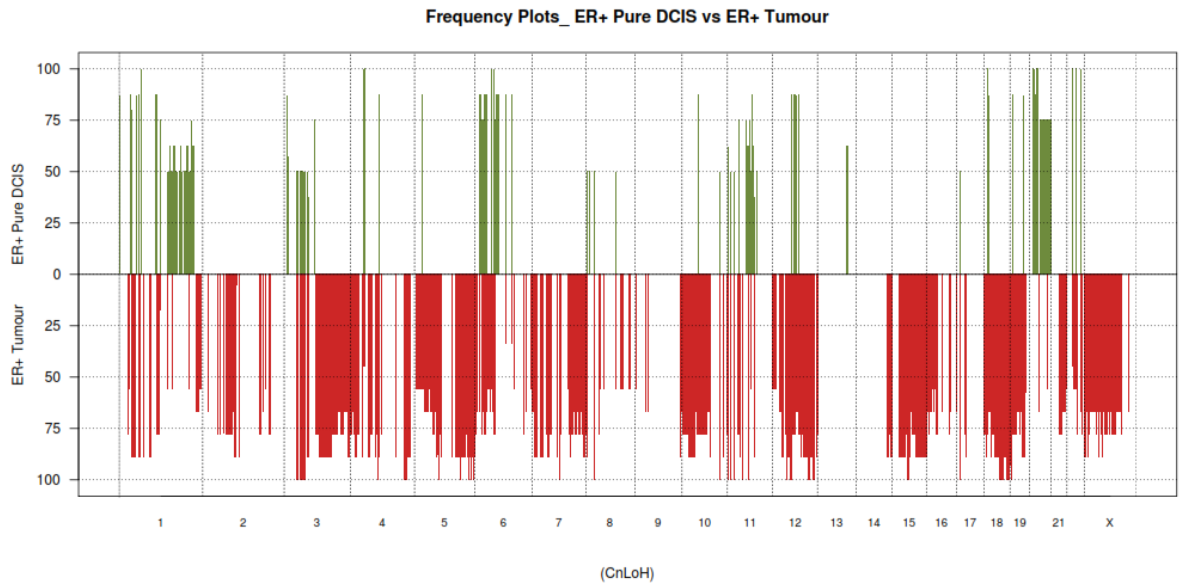


Frequency Plots_ ER+ Pure DCIS vs ER+ Tumour



Frequency Plots_ ER+ Pure DCIS vs ER+ Tumour





5.8.5.1 Amplification in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

There were no amplifications in ER positive pure DCIS (n=0/8) in this series. However, there were amplifications in ER positive invasive breast disease (n=5/9) not present in the ER positive pure DCIS.

1. p-values < 0.05;
2. A gene list is mapped from 2 genomic regions on chromosome 8;
3. These regions encompass 3 gene altered in ER positive invasive breast disease;

Genes:

AC103863.1 AF130342.1 EIF3H

PANTHER analysis: 1 mapped id is found, 2 mapped ids are not found.

Gene Information: ENSG00000147677 Protein ID: O15372

Gene Name: Eukaryotic translation initiation factor 3 subunit H Gene Symbol(s): EIF3H

Organism: Homo sapiens Alternate Ids: EIF3H_HUMAN (UniProtKB-ID) NP_003747 (refSeq)

ENSG00000147677 (Ensembl) 4503515(GI) IPI00977658 (IPI) EIF3H (Synonym)
 603912(MIM) ENST00000521861(Ensembl_TRS) HGNC:3273(HGNC)
 ENSP00000429931(Ensembl_PRO)

5.8.5.2 Duplication in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

No duplications were found in either ER positive pure DCIS or ER positive invasive breast disease.

5.8.5.3 Genomic Gains in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

Gains present in ER positive pure DCIS (n=1/8) were present in gains found in ER positive invasive breast disease (n=7/9). This represents gains in both ER positive pure DCIS and ER positive invasive breast disease.

1. p-values < 0.05;
2. A gene list is mapped from 16 genomic regions found on chromosomes 6, 16, 20;
3. These regions encompass 91 genes altered in ER positive invasive breast disease and ER positive DCIS;

Genes:

AC006076.1	AC106739.1	AL109840.1	AL121583.1	AL121673.1	AL136532.1	ARF4P2	ASXL1	ATP5E	C20orf11
C20orf112	C20orf166	C20orf195	C20orf20	C20orf200	C20orf90	C6orf48	COL9A3	CTD-3184A7.4	CTS2
DIDO1	DNMT3B	DPH3B	EEF1A2	GATA5	GCNT2	GMEB2	GNAS	GNASAS	HLA-DQA1
DRB1	hsa-mir-1-1	hsa-mir-133a-2	HSPA1A	HSPA1B	HSPA1L	IL4R	JMJD5	KCNQ2	KIF3B
NSMCE1	NTSR1	OGFR	PPDPF	PTK6	RP11-11M20.2	RP11-358D14.2		RP11-360O19.1	
RP11-360O19.4		RP11-93B14.4		RP11-93B14.5		RP11-93B14.6		RP1-1J6.2	RP1-290I10.3
RP1-290I10.4		RP1-290I10.5		RP1-290I10.6		RP1-290I10.7		RP1-290I10.8	RP1-
309F20.2	RP1-309F20.3	RP13-30A9.1		RP13-30A9.2		RP4-543J19.8		RP4-697K14.12	RP4-
697K14.3	RP4-697K14.7	RP4-724E16.2		RP5-1184F4.5		RP5-885L7.10		RP5-885L7.11	
RPL7P3	RTSL1	SLC17A9	SLCO4A1	SLMO2	SNTA1	SRMS	STMN3	TAF4	TCFL5
TH1L	TOP1	TUBB1	VARS	XXbac-BPG254F23.5		XXbac BPG254F23.6		ZNF217	

PANTHER analysis: 43 mapped ids are found, 48 mapped ids are not found.

There were gains present in ER positive invasive breast disease (n=6/9) not observed in ER positive pure DCIS (n=0/8).

1. p-values < 0.05;
2. A gene list is mapped from 12 genomic regions found on chromosomes 6, 18, 20:
3. These regions encompass 91 genes altered in ER positive invasive breast disease;

Genes:

5S_rRNA	7SK	AIM1	AL031963.1	AL031963.2	AL031963.3	AL031963.5	AL033383.2	AL033523.1	AL109754.1
AL133351.1	AL133509.1	AL138831.1	AL139092.1	AL160398.1	AL445309.1	AL445309.2	ANKRD5	AP005058.2	AP005482.1
AP006565.1	ATG5	BPHL	C6orf145	C6orf146	C6orf195	C6orf201	CDYL	CEP76	FAM136B
FAM50B	FOXO3	LACE1	MYLK4	NQO2	PECI	PRDM1	PRPF4B	PSMG2	PSMG4
RIPK1	RP11-145H9.3		RP11-404H14.1		RP1-140K8.1		RP1-140K8.2		RP1-140K8.3
RP1-140K8.4		RP11-416N4.1		RP11-416N4.4		RP11-420G6.3		RP11-420G6.4	
RP11-420L9.2		RP11-420L9.4		RP11-428J1.2		RP11-620A17.1		RP11-625P7.1	
RP11-697G4.3		RP1-223B1.1		RP1-40E16.11		RP1-40E16.2	RP1-40E16.3		RP1-40E16.4
RP1-40E16.8		RP1-40E16.9		RP1-90J20.10		RP1-90J20.2	RP1-90J20.7	RP1-90J20.8	RP3-359N14.1
RP3-359N14.2		RP3-400B16.1	RP3-400B16.2		RP3-400B16.3		RP3-406P24.1		RP3-406P24.3
RP3-430A16.1	RP3-527G5.1		RP4-529N6.1		RP5-839B4.7		RP5-839B4.8		SERPINB1
SERPINB6	SERPINB9	SLC22A23	SNAP25	snoU13	SPIRE1	TUBB2A	TUBB2BU6	WRNIP1	Y_RNA

PANTHER analysis: 26 mapped ids are found, 66 mapped ids are not found.

5.8.5.4 Genomic Sc Gains in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

Sc gains in ER positive pure DCIS (n=5/8) show some overlap with Sc gains in ER positive invasive breast disease (1/9).

1. p-values < 0.05;
2. A gene list is mapped from 6 genomic regions found on chromosomes 1, 8;
3. These regions encompass 28 genes altered in ER positive pure DCIS;

Genes:

ACBD6	AL359853.1	AL359853.3	AL390718.1	C1orf192	CEP350	FMO2	FMO3	FMO6P	hsa-mir-1295	LHX4
QSOX1	RP11-122G18.5		RP11-12M5.4		RP11-145A3.1		RP11-145A3.2		RP11-216N21.1	

RP1-127D3.4	RP11-502H18.2	RP11-521I8.1	RP11-533E19.2	RP11-533E19.3	RP1-45C12.1
RP5-1180C10.2	SDHC	TOR1AIP1	TOR1AIP2	U6	

PANTHER analysis: 11 mapped ids are found, 17 mapped ids are not found.

Sc gains in ER positive invasive breast disease (6/9) show some overlap with the Sc gains in ER positive pure DCIS (n=1/8).

1. p-values < 0.05;
2. A gene list is mapped from 10 genomic regions found on chromosomes 6, 20;
3. These regions encompass 67 genes altered in ER positive invasive breast disease;

Genes:

AL031963.1	AL031963.2	AL031963.5	AL109754.1	AL109840.1	AL121583.1	AL121673.1	AL133343.1	AL136532.1	ANKRD5	
ARF4P2	ASXL1	ATP5E								
BPHL	C20orf11	C20orf112	C20orf166	C20orf20	C20orf200	C20orf90	COL9A3	CTS2	DIDO1	DNMT3B
	DPH3B	GATA5	GNAS	GNASAS	hsa-mir-1-1	hsa-mir-133a-2		KIF3B	NTSR1	OGFR
	RP11-11M20.2	RP11-410N8.1		RP11-410N8.3		RP11-410N8.4		RP11-416N4.1		RP11-416N4.4
	RP11-93B14.4		RP11-93B14.5		RP11-93B14.6		RP1-1J6.2	RP1-309F20.2	RP1-309F20.3	
	RP13-30A9.1		RP13-30A9.2		RP1-40E16.11		RP1-40E16.8		RP1-40E16.9	RP4-
543J19.8	RP5-1184F4.5		RP5-839B4.7		RP5-839B4.8	RP5-885L7.10		RP5-885L7.11		RPL7P3
	SLC17A9	SLCO4A1	SLMO2	SNAP25	TAF4	TCFL5	TH1L	TOP1	TUBB1	TUBB2A

PANTHER analysis: 23 mapped ids are found, 44 mapped ids are not found.

5.8.5.5 Losses in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

Genomic losses in ER positive pure DCIS (n=6/8) show some overlap with genomic losses in ER positive invasive breast disease (1/9).

1. p-values < 0.05;
2. A gene list is mapped from 32 genomic regions on chromosome 11;
3. These regions encompass 225 genes altered in ER positive pure DCIS;

Genes:

5S_rRNA	7SK	AC009294.1	AC019227.1	AC019227.2	ACA59	ACAD8	ACRV1	ADAMTS15	ADAMTS8	
	AMICA1	ANKK1	AP000757.1	AP000770.1	AP000797.1	AP000797.2	AP000843.2	AP000844.1	AP000880.1	AP000892.4
AP000907.1	AP000907.2	AP000907.3	AP000908.1	AP000911.1	AP000925.2	AP000926.1	AP000936.1	AP000936.3	AP000936.4	
	AP000936.5	AP001122.1	AP001267.1	AP001267.2	AP001582.1	AP001781.1	AP001891.1	AP001979.1	AP001992.2	AP001998.1

5.8.5.6 Total Loss in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

There are total losses for ER positive pure DCIS (n=5/8) not observed in ER positive invasive breast disease (n=0/9).

1. p-values < 0.05;
2. A gene list is mapped from 12 genomic regions found on chromosomes 9, 11, 15, 19;
3. These regions encompass 22 genes altered in ER positive pure DCIS;

Genes:

7SK	AC025678.1	AC027139.2	AC027139.3	AP000720.1	AP001482.1	AP003072.1	CCDC67	CTSC	GOLGA8B	GRM5	hsa-
mir-1233	KDM4C	NOX4	OR10V1	OSBP	PATL1	RP11-657K20.1	SLC36A4	TYR	U7		
	ZNF208	ZNF257									

PANTHER analysis: 13 mapped ids are found, 10 mapped ids are not found.

There were total losses present for ER positive invasive breast disease (n=9/9) not observed in ER positive pure DCIS (n=0/8).

1. p-values < 0.05;
2. A gene list is mapped from 84 genomic regions found on chromosomes 8, 13, 14, 16, 17, X;
3. These regions encompass 443 genes altered in ER positive invasive breast disease;

Genes:

5S_rRNA	7SK	AC000003.1	AC001226.2	AC002347.1	AC002366.3	AC002504.1	AC005284.1	AC005323.1	AC005410.2
		AC005548.1	AC005548.2	AC005696.2	AC005696.3	AC005725.1	AC007333.1	AC007333.2	AC015799.1
									AC015799.2
									AC015908.2
		AC016292.3	AC022716.1	AC025518.1	AC026462.1	AC027045.1	AC027763.1	AC027763.2	AC073493.1
									AC073493.2
									AC073493.3
		AC074035.1	AC087498.1	AC090282.1	AC097370.1	AC099684.1	AC112778.1	AC112778.2	AC144838.1
									AIPL1
									AL136001.1
		AL136160.1	AL136219.1	AL136359.1	AL138690.1	AL138815.1	AL138815.3	AL139002.1	AL139034.1
									AL139082.1
									AL158032.2
AL162377.1		AL162377.2	AL163544.1	AL355611.1	AL355611.2	AL355611.3	AL356863.1	AL359180.1	AL391384.1
									AL442203.2
AL445989.1		AL450423.1	AL450447.1	AL512362.1	AL512652.1	AL589987.1	AL590007.2	AL596092.1	ALOX12
									AMAC1L3
ARHGAP6	ARX	ASGR1	ASGR2	ASPA	ASPG	ASS1P4	ATP7B	ATP8A2	ATXN8OS
									B3GALT
									BCL6B
BTF3L1	BX649443.1		BX890604.1	BX890604.2	C13orf23		C13orf34	C13orf36	C13orf37
									C14orf180
									C17orf100
C17orf49	C17orf51		C17orf61	C17orf74		C17orf91	CCDC42	CCDC70	CCNA1
									CDK8
									CHRNA1
									CHRNA1
CLN5	CRYL1		CTD-2545G14.1		CTNS		DACH1	DGKH	DHRS7C
									DHRSX
									DHX9P1
DIAPH3	DIS3		DLC1	DLG4	DNAH9		DPH1	DUSP21	EBF2
									EDNRB
									EFHA1
EIF4A1P6	FAM123A		FAM64A	FBXL3	FBXO39	FGF11	FGF9	FLT1	FOXO1
									FREM2
									FREQ

	GAPDHP34	GAS7	GFRA2	GJA3		GJB2	GJB6	GLP2R		GPR12	HDHD1A	HMCN2	HNFG4P1		hsa-mir-1253	hsa-
mir-195		hsa-mir-203	hsa-mir-22		hsa-mir-497	hsa-mir-744		hsa-mir-759	IFT88		IL17D	IRG1		KCNJ12		
	KDM6A	KIAA0564	KIAA0664		KIAA0753	KIF26A		KLF12	KLF5		KLHL1	LHFPLIN28AP2			MAP2K3	
	MAP2K4	MED31	METT10D		MFS6L	MTUS2	MXRA5	MYH1	MYH10		MYH13	MYH2		MYH3	MYH4	
	MYH8	N6AMT2	NDEL1	NEFL		NEFM		NEK3	NEK5		NFYAP1	NHLRC3		NLGN2		
	NLGN4X	NTN1	NXN		OLFM4	OR1A1		OR1A2	OR1D2	OR1D5	OR1E1	OR1E2		OR1G1		
	OR3A1	OR3A2	OR3A3	OR3A4	OVCA2	PABPC3		PAFAH1B1	PARP4		PCDH8	PIBF1		PIK3R5		
	PIK3R6	PIRT		PITPNM3		PLSCR3		POLA1	POLR2A	POU4F1	PRKX		PSPC1	RAP1GAP2		RCVRN
	RFXAP	RNASEK	RNF17		RNF219	RNF6	RP11-101P17.10		RP11-101P17.11	RP11-101P17.5				RP11-101P17.6		
RP11-101P17.7		RP11-101P17.8			RP11-101P17.9			RP11-107B1.1			RP11-110K18.2			RP11-110K8.1		
		RP11-117N4.1		RP11-11C5.1			RP11-12I24.2	RP11-12I24.3			RP11-133M24.1			RP11-137M6.2		
RP11-149B7.1		RP11-149B7.3			RP11-149B7.4			RP11-149B7.5			RP11-165I9.3			RP11-165I9.4		
RP11-165I9.6	RP11-165I9.8		RP11-169O17.1			RP11-169O17.5			RP11-16D22.1	RP11-16L6.3				RP11-172E9.2		
		RP11-172H24.3	RP11-181D10.2			RP11-187A9.2		RP11-187A9.3			RP11-196I2.1			RP11-196P2.1		
	RP11-197L7.2	RP11-1M18.1	RP11-203P2.1			RP11-203P2.2			RP11-206H15.2		RP11-206L1.2					
	RP11-209P2.1		RP11-214O11.1			RP11-214O11.2			RP11-218I21.1		RP11-248G5.3					
	RP11-24H2.1		RP11-24H2.2			RP11-252B20.1	RP11-261P13.3		RP11-264J4.4		RP11-264J4.5					
	RP11-264J4.6		RP11-266E6.2			RP11-267I18.1		RP11-271B5.1		RP11-279N8.1	RP11-304M3.2					
	RP11-30C8.1		RP11-30C8.2			RP11-318G21.2		RP11-318G21.3		RP11-318G21.4		RP11-				
323P14.2	RP11-325D5.3		RP11-327P2.3			RP11-335N6.1		RP11-342C20.2		RP11-342C20.3						
	RP11-349O10.1		RP11-350A18.1			RP11-351N4.3		RP11-360A9.1		RP11-364B14.1						
	RP11-364B14.2		RP11-364B14.3			RP11-367C11.2	RP11-380N8.1		RP11-380N8.3		RP11-380N8.4					
	RP11-381L18.2		RP11-388E20.1			RP11-393H6.2		RP11-393H6.3		RP11-394A14.1						
	RP11-394A14.2		RP11-394A14.3			RP11-398O19.3		RP11-408O13.2		RP11-408O13.3		RP11-				
421P11.5		RP11-430I3.2			RP11-430K22.1			RP11-430K22.2			RP11-431O22.2			RP11-442F12.2		
	RP11-459A10.1		RP11-459A10.2			RP11-459A10.3			RP11-459J23.1		RP11-459J23.2					
	RP11-462C21.1		RP11-463M3.1	RP11-467D10.2			RP11-471M2.1		RP11-471M2.2		RP11-473M10.1					
	RP11-473M10.2		RP11-473M10.3			RP11-474L7.4		RP11-483M24.2		RP11-501G6.1	RP11-505F3.2					
	RP11-505F3.4		RP11-518D7.1			RP11-51B13.2		RP11-520F24.1	RP11-520F24.2		RP11-520F24.3					
	RP11-520F9.2		RP11-521H3.1			RP11-523H24.5		RP11-527F15.1		RP11-52L5.4	RP11-534K14.1					
	RP11-534K14.2	RP11-545M8.1		RP11-545M8.2			RP11-545M8.3		RP11-556N21.4	RP11-56M2.1	RP11-570F6.1					
	RP11-571G1.1		RP11-571G1.2			RP11-609D21.1	RP11-61K9.2		RP11-629E24.2		RP11-629O4.1					
RP11-706O15.1		RP11-706O15.3			RP11-706O15.5			RP11-706O15.7		RP11-706O15.8		RP11-707P20.1				
	RP11-713H12.1		RP11-76N11.2			RP11-77P3.1			RP11-7B3.2	RP11-7B3.3	RP11-7B3.4		RP11-822E23.1			
	RP11-88G17.6		RP11-90M5.1	RP11-963H4.1			RP1-258N20.3		RP13-858C7.1		RP13-926M18.1					
RPA1	RTN4RL1	SCEL	SCO1		SERPINF1	SERPINF2		SHISA6	SHPK		SLAIN1	SLC13A5		SLC16A11		
	SLC16A13	SLITRK5	SMAD9		SMYD4	SNORA25		SNORA48	SNORA68		SNORA9	SNORD116		SNORD37		
snoU13		SPATA22	SPDYE4		SPEM1	STOML3		STX8	TCEB1P23	TDRD3	TDRD9	TEKT1		TLCD2		
	TMEM102	TMEM179	TMEM46	TNK1	TOX3	TRPC4		TRPV1	TRPV3		TXNDC17	U2		U3		U4
U6	U7	UFM1	UTP14C		VCX3A	WASF3		WDR81	WSCD1		XAF1	XPO4	Y_RNA		ZBED1	
	ZBTB4		ZDHHC20	ZMYM2	ZMYM5											

PANTHER analysis: 167 mapped ids are found, 276 mapped ids are not found.

5.8.5.7 CdLOH in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

There are similarities between CdLOH in ER positive pure DCIS (n=7/8) and CdLOH in ER positive invasive breast disease (n=3/9).

1. p-values < 0.05;
2. A gene list is mapped from 38 genomic regions found on chromosomes 9, 11, 17;
3. These regions encompass 281 genes altered in ER positive invasive breast disease and ER positive pure DCIS;

Genes:

5S_rRNA	7SK	AC004148.1	AC006435.1	AC009294.1	AC019227.1	AC019227.2	AC055839.2	AC087500.1	AC087742.1			
		AC116914.1	AC129492.6	ACA59	ACAD8	ACRV1	ADAMTS15	ADAMTS8	ALOX12B	ALOX15B	ALOXE3	AMICA1
ANKFY1	ANKK1	AP000757.1	AP000770.1	AP000797.1	AP000797.2	AP000843.2	AP000844.1	AP000880.1	AP000892.4			
		AP000907.1	AP000907.2	AP000907.3	AP000908.1	AP000911.1	AP000925.2	AP000926.1	AP000936.1	AP000936.3	AP000936.4	
		AP000936.5	AP001122.1	AP001267.1	AP001267.2	AP001582.1	AP001781.1	AP001891.1	AP001979.1	AP001992.2	AP001998.1	
AP002451.1		AP002806.1	AP002806.2	AP002856.2	AP002856.4	AP002856.5	AP002856.6	AP002856.7	AP002884.4	AP002886.2	AP002962.1	
		AP002986.1	AP002991.1	AP003025.2	AP003039.3	AP003039.4	AP003040.1	AP003040.2	AP003041.1	AP003070.1	AP003402.1	
		AP003402.2	AP003500.1	AP003730.1	AP003730.2	AP004248.4	AP004372.1	AP005597.1	AP005638.1	AP005638.3	AP005639.1	
AP006216.10	AP006216.11		AP006216.12		AP006216.5	APLP2	APOA1	APOA4		APOA5		
APOC3	ARHGAP20	ARHGAP32	ARHGEF17	ATP2A3	ATP5L	AURKB	B3GAT1	BACE1	BAK1P2	BARX2		
BCO2	BSX	BTG4	BUD13	C11orf1	C11orf44	C11orf45	C11orf52	C11orf57	C11orf63		C11orf71	
C11orf88	C17orf44	C17orf59	C17orf68	C17orf85	C17orf87		C1QBP	CADM1	CAMKK1	CD3D		
CD3E	CD3G	CEP164	CHEK1	CLDN25	CNTROB	CRTAM	CRYAB	CTNS	CUL5	CYB5D2		
DDX10	DDX25	DDX6	DERL2	DHX33	DIXDC1	DLAT	DRD2	DSCAML1		EI24		
ETS1	FAM168A	FAM55A	FAM55B	FAM55D	FDX1	FDXACB1	FLI1		FXSD2	FXSD6		
GLB1L2	GLB1L3	GSG2	GUCY2D	HES7	HIC1	hsa-mir-132		hsa-mir-212	hsa-mir-34b	hsa-mir-34c		
HSPA8	HSPB2	HTR3A	HTR3B	HYLS1	IGSF9B	IL10RA	IL18	ITGAE	JAM3	KCNJ1		
KCNJ5	LAYN		METT10D	MIS12	MLL	MPZL2	MPZL3	NCAM1	NCAPD3	NCRNA00167	NEU3	
NFRKB	NLRP1	NNMT	NUP88	OPCML	OR4A15	OR4A16	P2RX1	P2RX5	PAFAH1B2	PATE1		
PATE2	PATE3	PATE4	PCSK7	PER1	PFAS	PIH1D2	POU2AF1	PPIHP1	PPP2R1B	PRDM10		
	PTPRD	PUS3		RAB39	RABEP1	RANGRF	RBM7	RDX	RELT	REXO2	RNF214	RP11-
237N19.1		RP11-48B14.1		RP11-567M21.1		RP11-667M19.1	RP11-762B21.1		RP11-762B21.2			
		RP11-762B21.3		RP11-801G16.1		RP11-832N8.1		RPAIN	RPL37AP8	SCARNA11	SCN2B	
SCN4B	SDHD	SIDT2	SIK2	SIK3	SLC25A35	SLC35F2	SMG6	SNORD14	snoU13	SNX19	SPATA19	
SPCS2	ST14	STT3A	TAGLN	TAX1BP3	TEX12	THYN1	TIMM8B		TMEM107	TMEM225		
TMEM45B	TMEM93	TMPRSS13	TMPRSS4	TMPRSS5	TP53AIP1	TRIM48	TTC12	U2	U4			
U4atac	U6	U6atac	U7	U8	UBASH3B	UBE4A	USP28	USP6	VAMP2	VPS26B	XRRA1	
Y_RNA	ZBTB16	ZBTB44	ZC3H12C	ZNF259	ZNF594	ZW10	ZZEF1					

PANTHER analysis: 168 mapped ids are found, 113 mapped ids are not found.

There are some similarities between CdLOH in ER positive invasive breast disease (n=7/9) and CdLOH in ER positive pure DCIS (n=1/8).

1. p-values < 0.05;
2. A gene list is mapped from 34 genomic regions found on chromosomes 1, 8, 9, 15, 16, 19, 22, X;
3. These regions encompass 51 genes altered in ER positive invasive breast disease and ER positive pure DCIS;

Genes:

5S_rRNA	7SK	AC011445.1	AC018558.1	AC019176.2	AC022716.1	AC023824.2	AC073493.1	AC073493.2	AC073493.3
AF238378.1	AL611925.1	AP000351.4	BMP15	BX649443.1	CACNA1B	CSMD1	DDT	DEFA5	DHRX
EHMT1	FRMPD4	GABRB3	GANC	GFRA2	GSTT2	GSTTP1	hsa-mir-627	IL28A	IL28B
LARGE	MAGED1	NCCRP1	PLA2G4F	PRKCZ	RP11-104D21.1		RP11-104D21.2		RP11-104D21.3
RP11-637B23.1		RP11-637B23.3		RP13-346H10.2		RP13-858C7.1		RP5-1168A5.1	RP5-
1170D6.1	RP5-964N17.1		SNORA70	SYCN	TMEM87A	U3	VPS39		

PANTHER analysis: 20 mapped ids are found, 31 mapped ids are not found.

5.8.5.8 CnLOH in Oestrogen Receptor Positive Pure DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

There are similarities between ER positive pure DCIS (n=8/8) and ER positive invasive breast disease (n=4/9).

1. p-values < 0.05;
2. A gene list is mapped from 67 genomic regions found on chromosomes 1, 3, 4, 5, 6, 10, 11, 12, 18, 19, 20, 22;
3. These regions encompass 357 genes altered in ER positive pure DCIS;

Genes:

5S_rRNA	7SK	ABT1	AC000032.1	AC002485.1	AC005399.1	AC025423.2	AC034193.1	AC099680.1	AC124890.1
RP11-209A2.1		ADH5P4	AGAP10	AHCYL1	AK3L1	AKD1	AL021807.2	AL021808.1	AL021808.2
							AL021808.3		
AL023913.1	AL035045.1	AL035661.3	AL050321.1	AL050321.2	AL050322.1	AL050335.1	AL050335.2	AL109754.1	AL109922.1
AL109947.1	AL109947.2	AL118509.1	AL121584.1	AL121988.1	AL132996.1	AL133255.1	AL133472.1	AL138850.1	AL139429.1
AL354680.1	AL391137.2	AL391559.1	AL589723.1	AL590062.1	AL591044.1	AL591044.2	AL591044.3	AL591044.4	AL691432.2
AL713998.1	ANKRD5	ANXA8L1	AP001029.1	AP001992.2	AP005482.1	APBB2	ASPHD2	BACH2	BANF2
BEND6	BMS1P2	BTBD3	BX255972.1	C1orf212		C20orf12	C20orf187	C20orf3	C20orf61
									C20orf72
C20orf78	C20orf79	C20orf82	C20orf94	C3orf10	C6orf1	C6orf140	C6orf184	CD164	CDH12
									CDK11A
CECR2	CECR9	CENPM	CEP76	CRYBB2P1	CSF1	CSRP2BP	CST1	CST2	CST3
								CST4	

CST5	CST7	CSTP1	CSTP2	CTA-221G9.9	CTA-246H3.11	CTA-246H3.7			
CTA-246H3.8		CTA-250D10.15	CTA-250D10.19	CTB-1048E9.7	CTSLL7	CYP2S1	DEK		
DHFRP2	DLGAP3	DTNBP1	ELAVL4	ENSAP1	EPS8L3	EYS	FAM25B	FAM35B2	FAM40A
FAT1P1	FGFR3P	FIG4	FLOT1	GAPDP2	GGTLA4	GJA4	GJB3	GJB4	GJB5
							GLS2		GLUDP8
GMPR	GSTM2	GSTM3	GSTM3P	GSTM5	GUSBL1	HCP5	HIST1H2AH	HIST1H2BJ	HIST1H2BK
									HIST1H4I
HLA-S	HMGA1	HNRNPM	HPS4	hsa-mir-33a	IRAK2	JARID2	KDM1B	KHDRBS2	KIF13A
									KIF16B
									LGSN
LIN28AP3	LRP5L	MACROD2	MAP3K7	MAR02	MDC1	MDM2	MICA	MICAL1	MKKS
MMP23A	MRPS11P1	MYLIP	MYO1F	N4BP2	NAV3	NCRNA00227		NCRNA00240	NEU3
									NHLRC1
NKX2-2	NKX2-4	NUDT3	NXT1	OTOR	OVOL2	PA2G4P2	PACIN1	PAK7	PGAM3P
									PHACTR1
PLK1S1	POM121L1P	POM121L2	POM121L3P	PPIAP17	PPIL6	PPYR1	PRAM1	PRSS16	PSMG2
PTMAP3	PTPN2	RBMS2	RNF11B	RNF144B	RP11-102J14.1	RP11-124A7.2		RP11-129K24.4	
RP11-144G6.10		RP11-144G6.4		RP11-144G6.9		RP11-146I2.1		RP11-157E14.1	
RP11-157E14.3	RP11-195M16.1		RP11-218C14.5		RP11-227D2.3	RP11-239L20.3		RP11-239L20.4	RP11-239L20.5
RP11-292F22.3		RP11-292F22.4		RP11-292F22.5		RP1-129L7.1	RP11-307F22.2		RP11-314P12.2
RP11-314P12.3	RP11-388K2.1		RP1-13D10.2		RP1-13D10.3		RP1-13D10.4		RP1-13D10.5
	RP11-401E14.1	RP11-401E14.2	RP11-416N4.1		RP11-416N4.4		RP11-457M11.2		RP11-492I2.1
	RP11-513I15.3		RP11-513I15.5	RP11-513I15.6		RP11-526K17.2	RP11-528A10.1		RP11-528A10.2
RP11-542F9.1	RP11-560J1.1		RP11-575L16.2	RP11-69I22.2		RP11-69I22.3		RP11-707M13.1	
	RP11-74E24.2		RP1-187N21.2		RP1-190J20.2		RP1-198K11.3	RP1-240B8.3	RP1-298M8.1
	RP3-322G13.7		RP3-453I5.2	RP3-462D8.2		RP3-525L6.2	RP4-564O4.1		RP4-568F9.3
									RP4-568F9.6
568F9.7	RP4-580G13.1		RP4-697P8.2		RP4-697P8.3		RP4-705D16.3		RP4-717M23.2
									RP4-718D20.3
	RP4-726N1.2		RP4-734C18.1		RP4-735C1.4		RP4-735C1.5		RP4-735C1.6
									RP4-
737E23.2	RP4-737E23.4		RP4-742J24.2		RP4-777D9.2	RP4-794H19.4		RP4-796I8.1	RP5-1018A4.3
									RP5-
1068E13.3	RP5-1068F16.3		RP5-1069C8.1		RP5-1069C8.2		RP5-1077I2.3		RP5-1077I2.4
									RP5-1100I6.1
	RP5-1100I6.2		RP5-1177M21.1		RP5-1195D24.1	RP5-839B4.7		RP5-839B4.8	RP5-860P4.2
									RP5-
RP5-872K7.2		RP5-872K7.7		RP5-905G11.3	RP5-919F19.5	RP5-931K24.1		RP5-984P4.1	RP5-
984P4.4		RP5-997D16.2		RPL13P	RPL15P1	RPL23AP6	RPL24P2	RPL7AL3	RPLP0P1
									RPS10
RPS10P2	RPS11P1	RPS15AP1	RRBP1	SEP03	SEZ6L	SHISA8	SLC24A3	SMCHD1	SMPD2
									SNAP25
SNORD17	SNORD83	snoU13	SNRPB2	SNX5	SPCS2	SPDEF	SPIRE1	SPRYD4	SPTLC3
SREBF2	TASP1	TMEM90B	TNFRSF13C	TPMT	TUBB	U1	U2	U3	U6
									U7
									VHL
VN1R11P	VN1R12P	VN1R13P	VN1R14P	XRN2	XRRA1	XXbac-BPG181B23.4		XXbac-BPG181B23.6	XXbac-
BPG248L24.11	XXbac-BPG248L24.12		XXbac-BPG27H4.7	XXbac-BPG308K3.6				XXyac-YR14BB7.1	XXyac-
YX60D10.1	Y_RNA	Z99716.2	ZBTB24	ZDHHC20P2	ZNF133	ZNF322A	ZNF414		

PANTHER analysis: 120 mapped ids are found, 237 mapped ids are not found.

There are similarities between ER positive invasive breast disease (n=4/9) and ER positive pure DCIS (n=8/8).

1. p-values < 0.05;
2. A gene list is mapped from 34 genomic regions found on chromosomes 3, 4, 5, 7, 8, 10, 11, 12, 15, 17, 18, 19;
3. These regions encompass 470 genes altered in ER positive invasive breast disease;

Genes:

5S_rRNA	AAAS	ABHD14A	ABHD14B	AC002347.1	AC002481.5	AC005323.1	AC008021.1	AC008590.1	AC009271.1
AC009704.1	AC009779.1	AC010297.1	AC011260.2	AC011385.1	AC011385.3	AC012379.1	AC012379.2	AC013558.1	AC015689.1
AC021066.1	AC022031.1	AC023055.2	AC023509.1	AC025165.2	AC025165.7	AC025165.8	AC026689.1	AC040963.1	AC044839.1
AC044839.2	AC055715.1	AC068473.1	AC068888.1	AC068987.1	AC068988.1	AC069209.1	AC073573.1	AC078778.1	AC078864.1
AC079600.1	AC087893.3	AC090398.1	AC090666.1	AC091103.1	AC091551.1	AC091934.1	AC091934.2	AC097359.1	AC097359.2
AC097359.3	AC099557.1	AC100783.1	AC103681.1	AC103681.2	AC104186.1	AC104448.1	AC104850.1	AC105105.1	AC105316.1
AC107016.1	AC107016.2	AC112205.1	AC112215.2	AC112215.4	AC116904.1	AC116917.2	AC126118.1	AC126118.2	AC132217.4
AC156455.1	AC190387.1	ACACB	ACVR1B	ACVRL1	ACY1	ALAS1	ALDOAP1	ALKBH2	ALPK2
AMHR2	ANAPC5	ANKRD33	AP005273.1	AP006288.1	ARHGEF3	ATCAY	ATF7	ATP5G2	ATP9B
ATP9B	ATXN2	ATXN7	B4G	ALNT1	BAP1	BCL7A	BIN2	BLOC1S1	BMP3
C11orf94	C12orf10	C12orf44	C18orf26	C18orf54	C18orf55	C3orf45	C3orf49	C3orf63	C3orf78
C4orf22	C4orf39	CACNA2D2	CACNA2D3	CAMKK2	CAMP	CBX5	CCDC51	CCDC66	CCDC72
CCDC99	CCNG1	CCR8	CD63	CDC25A	CELA1	CHST11	CNDP1	CNDP2	COPZ1
CPE	CRY2	CSAD	CSRNP1	CSRNP2	CTB-164L20.1	CTB-95B16.1	CTD-3211M3.1	CTDP1	CTNNA1
CUX2	CX3CR1	CYB561D2	CYB5A	DAO	DAPK3	DAZAP2	DCD	DCTN2	DNAH1
DNAJC14	DOCK2	DTX3	DUSP7	DYM	EEF1A1P24	EEF2	EIF1B	EIF4B	EIF4E1B
ENTPD3	ESPL1	ETF1	FAF2	FAM109A	FAM196B	FAM69C	FBLL1	FBXW12	FGFR4
FHIT	GABRG2	GALNT6	GAS7	GDF11	GK3P	GLP2R	GLYCAM1	GLYCTK	GNAI2
GOLGA4	GPR84	GPRIN1	GRASP	GRP	GTSF1	GYLTL1B	HIGD1C	HIGD1DP	HMMR
HNRNPA1	hsa-let-7g	hsa-mir-103-1	hsa-mir-122	hsa-mir-135a-1	hsa-mir-148b	hsa-mir-1979	hsa-mir-2115	hsa-mir-218-2	hsa-mir-483
hsa-mir-578	hsa-mir-637	HSPA9	HYAL1	HYAL2	HYAL3	IFRD2	IGF2	IGF2AS	IGFBP6
IL31	INS	ITGA5	ITGA7	ITGB1BP3	ITGB7	KDM2B	KIF5A	KLHL2	KRT1
KRT2	KRT3	KRT4	KRT5	KRT6A	KRT6B	KRT6C	KRT7	KRT71	KRT72
KRT73	KRT74	KRT75	KRT76	KRT77	KRT78	KRT79	KRT8	KRT80	KRT81
KRT82	KRT83	KRT84	KRT85	KRT86	LACRT	LETMD1	LMO1	LRRC43	LRRFIP2
MAP2K2	MAP3K12	MAPK8IP1	MAT2B	MBD2	METTL7B	MEX3C	MFSD5	MIP	MLH1
MLXIP	MMP19	MXI1	MYH1	MYH13	MYH2	MYH4	MYH8	MYRIP	NACA3P
NAT6	NCKAP1L	NDUFB1P1	NEDD4L	NEFL	NEFM	NFATC1	NFE2	NISCH	NME6
NPFF	NR4A1	NSD1	NT5DC2	NUDCD2	OR10P1	ORMDL2	P2RX4	P2RX7	PANK3
PBRM1	PCBP2	PCBP4	PCDH24	PDE1B	PDZRN4	PEX16	PFDN5	PHF21A	PHF7
PIAS4	PLXNB1	POLI	POU6F1	PPM1M	PPP1R1A	PPTC7	PRDM11	PRICKLE1	PRKG2
PRR13	PSMD9	RAD9B	RAPGEF2	RARG	RARS	RASSF1	RCVRN	RDH5	RNF34
RNF44	RP11-100N20.1	RP11-1029M24.1	RP11-1029M24.2	RP11-138A23.1	RP11-138A23.2	RP11-153F5.1	RP11-155D18.11	RP11-168J18.4	RP11-168J18.6
RP11-219C20.1	RP11-241K18.1	RP11-241K18.2	RP11-24C3.2	RP11-25I15.1	RP11-285B24.1	RP11-294O2.2	RP11-330H6.5	RP11-331G2.4	RP11-366M4.1
RP11-366M4.11	RP11-366M4.12	RP11-366M4.13	RP11-366M4.14	RP11-366M4.15	RP11-366M4.17	RP11-366M4.2	RP11-366M4.3	RP11-366M4.4	RP11-366M4.5
RP11-366M4.6	RP11-366M4.8	RP11-391M1.2	RP11-391M1.3	RP11-416A17.1	RP11-438N5.1	RP11-462C21.1	RP11-469O11.2	RP11-502L5.2	RP11-504H5.1
RP11-50F24.3	RP11-51O17.1	RP11-528N21.2	RP11-528N21.3	RP11-541P9.3	RP11-554O1.1	RP11-644A7.1	RP11-727C1.1	RP11-834C11.3	RP11-834C11.4
RP11-834C11.5	RP11-834C11.6	RP11-834C11.7	RP11-834C11.8	RP11-972K6.1	RP5-1157M23.2	RPL14	RPL29	RPSA	RRM1
SARNP	SC4MOL	SCN8A	SEC11C	SEMA3B	SEMA3E	SEMA3G	SH2B3	SHC4	SIL1
SLC11A2	SLC25A38	SLC26A10	SLC35C1	SLC38A3	SLC4A8	SLIT3	SMAD4	SMAGP	SMNDC1
SMUG1	SNCB	SNORA30	SNORA62	SNORA64	SNORD28	SNORD37	SNORD63	snoU13	SOAT2
SOCS6	SP1	SP7	SPATA12	SPINK8	SPRYD3	STAB1	STARD6	STIM1	SVOP
SYT13	TARBP2	TCF4	TCTN1	TENC1	TFCP2	THOC7	TIMELESS	TLR9	TMEM115
TMEM192	TNNC1	TP53I11	TRIM60	TRIM61	TRIM75	TSPAN18	TTC21A	TUSC2	TUSC4
TWF2	TXNL1	U1	U4	U5					

U6	U6atac	U7	U73166.1	U73166.2	U73167.7	U73169.1	UNG	USP30	WDR51A	
WDR66	WDR7	WDR82	WWC1	XIRP1	Y_RNA	ZBTB7A	ZMYND10	ZNF346	ZNF385A	ZNF407
ZNF589	ZNF619	ZNF620	ZNF621	ZNF740						

PANTHER analysis: 285 mapped ids are found, 185 mapped ids are not found.

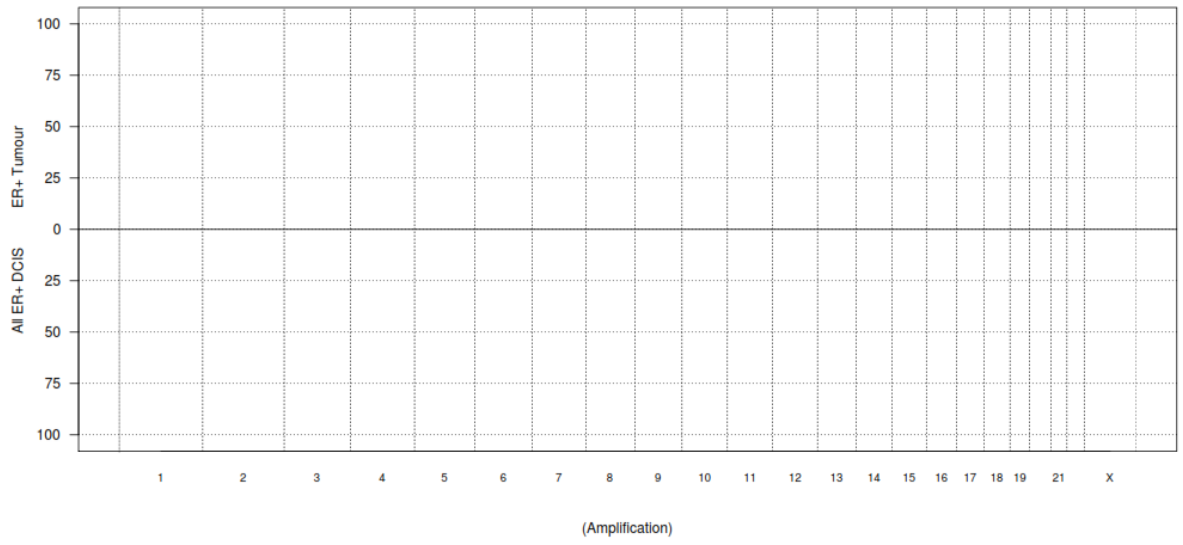
5.8.6 *Copy Number Aberrations for All Oestrogen Receptor Positive DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease*

These series examines the difference between all oestrogen receptor positive DCIS (i.e. both pure and that associated with invasion) and oestrogen receptor positive invasive breast disease.

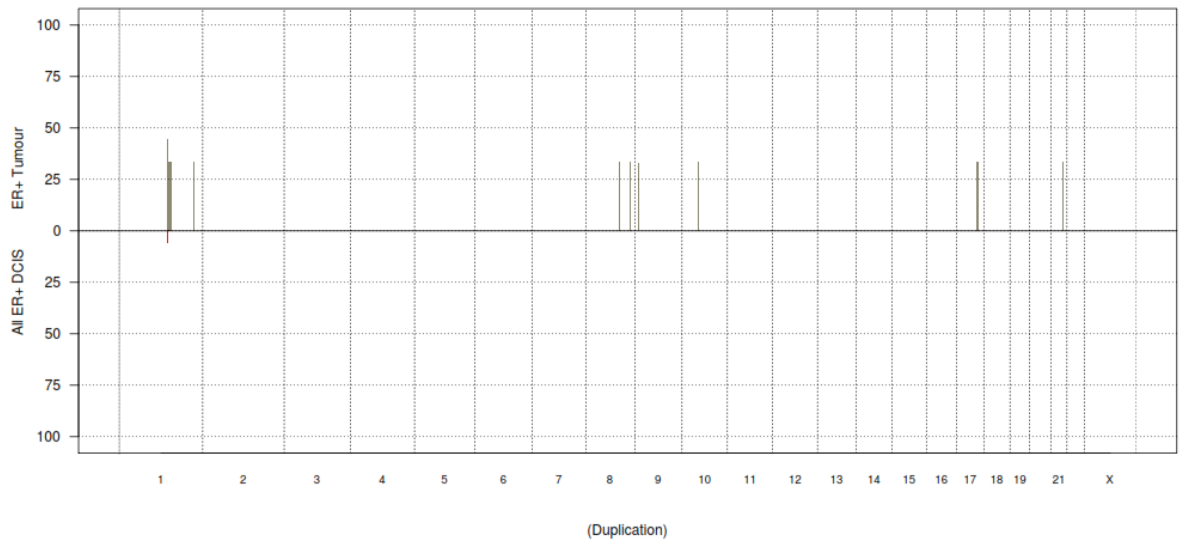
Frequency plots showing copy number aberrations between oestrogen receptor positive DCIS and oestrogen receptor positive invasive breast disease were provided by Breakthrough Breast Cancer/Research Oncology, King's College London Bioinformatics Department (Figure 39).

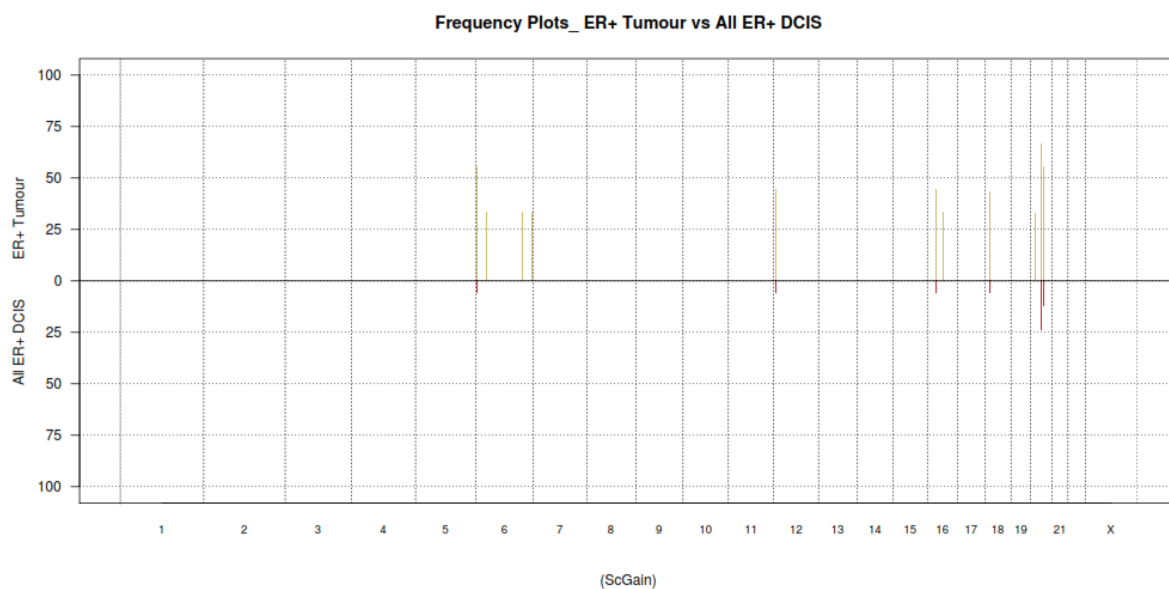
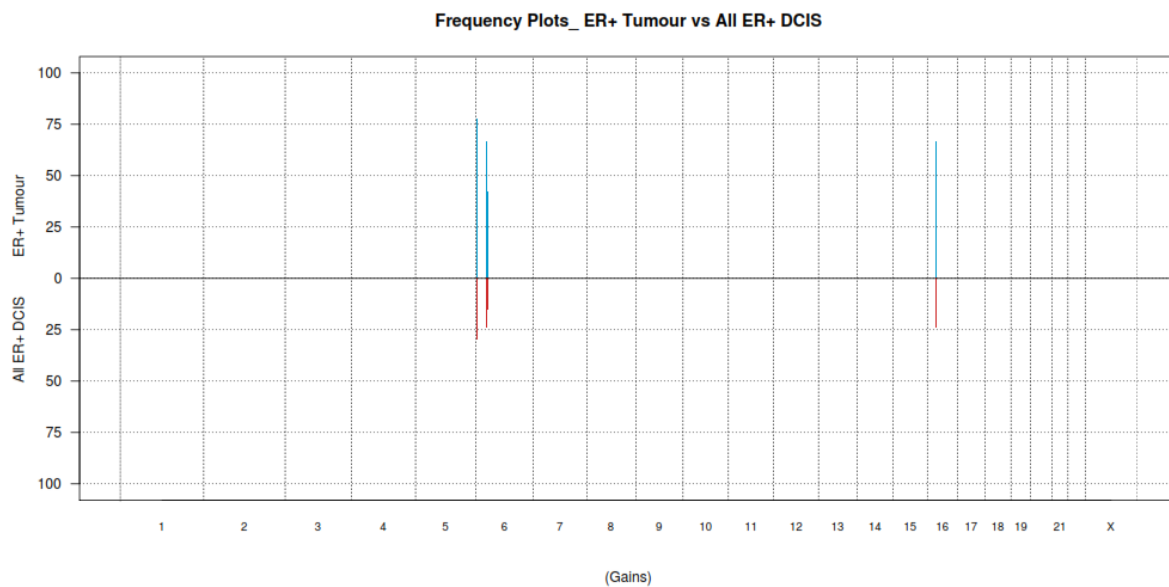
Figure 39: Frequency plots showing copy number aberrations between oestrogen positive DCIS and oestrogen positive invasive breast disease (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 209-212).

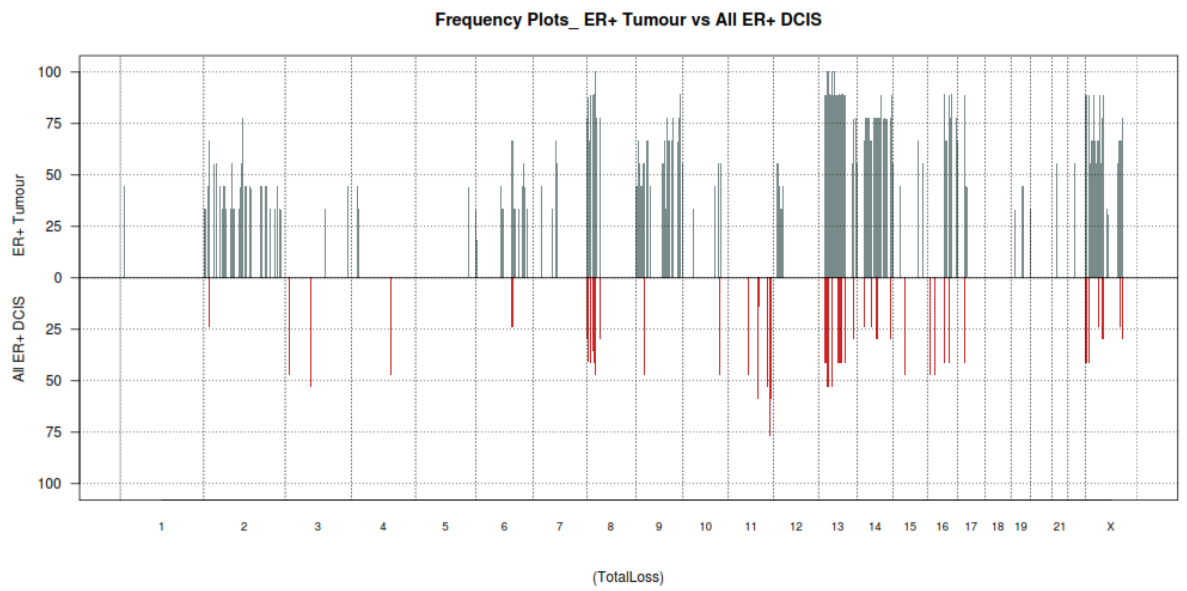
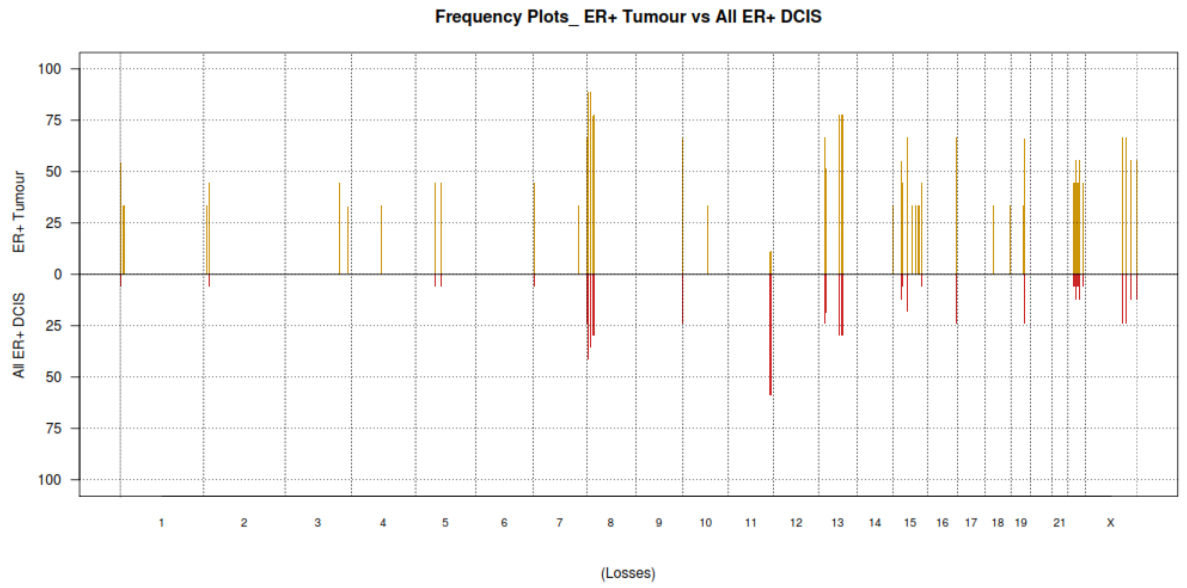
Frequency Plots_ ER+ Tumour vs All ER+ DCIS

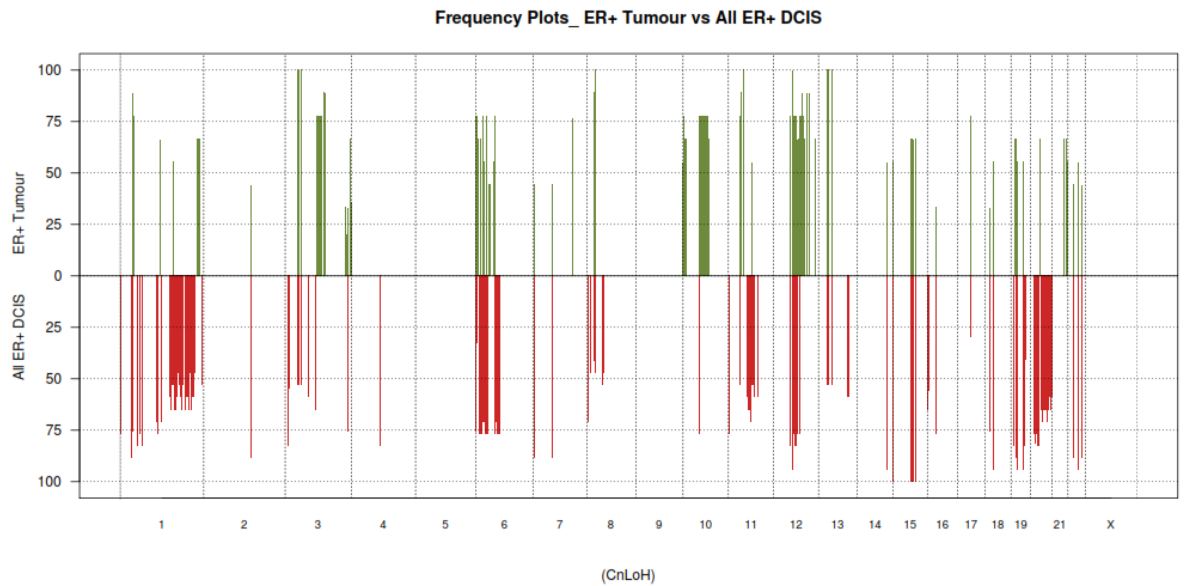
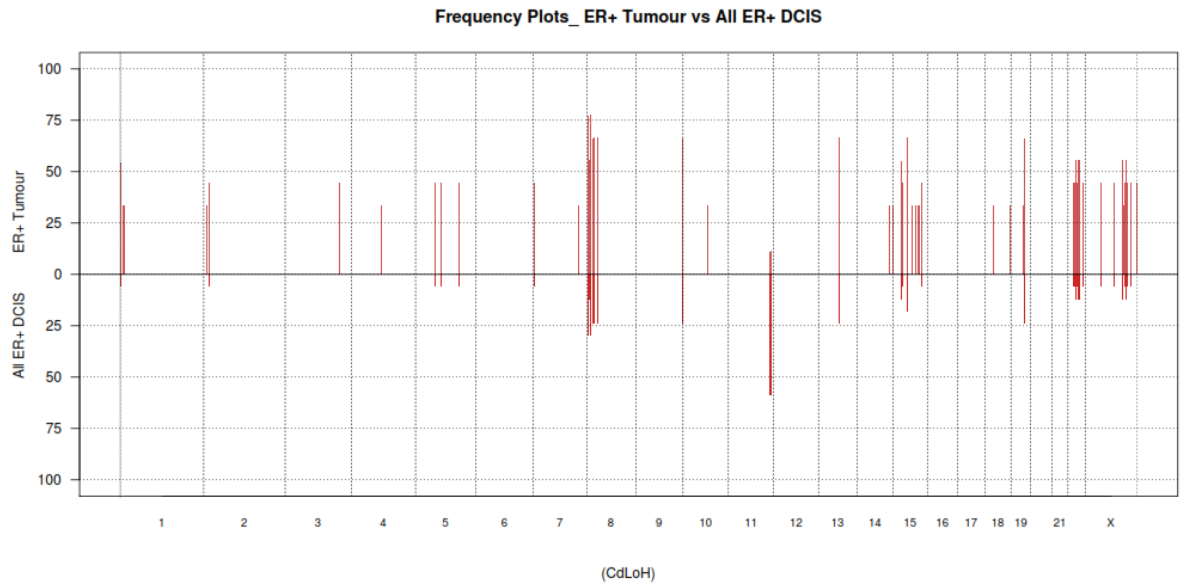


Frequency Plots_ ER+ Tumour vs All ER+ DCIS









5.8.6.1 Amplification in Oestrogen Receptor Positive DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

No amplifications in either ER positive DCIS or ER positive invasive breast disease are found in this series.

5.8.6.2 Duplication in Oestrogen Receptor Positive DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

There are some overlaps in the duplications present in ER positive DCIS (n=4/17) and in ER positive invasive breast disease (n=1/9).

1. p-values < 0.05;
2. A gene list is mapped from 2 genomic regions on chromosome 1;
3. These regions encompass 18 genes altered in ER positive DCIS and some ER positive invasive breast disease;

Genes:

AL592307.2	HFE2	TXNIP	POLR3GL	ANKRD34A	LIX1L	RBM8A	GNRHR2	PEX11B	ITGA10
ANKRD35	PIAS3	RP11-458D21.2		RP11-458D21.1	RP11-315I20.1		RP11-315I20.3		U1
PDE4DIP									

PANTHER analysis: 12 mapped ids are found, 6 mapped ids are not found.

There are duplications found in ER positive invasive disease (n=4/9) not present in ER positive DCIS (n=0/17).

1. p-values < 0.05;
2. A gene list is mapped from 17 genomic regions found on chromosomes 1, 8, 9, 10, 17, 21;
3. These regions encompass 255 genes altered in ER positive invasive breast disease;

Genes:

ADAM15	5S_rRNA	7SK	AC046185.1	AC046185.2	AC110921.1	ADAR	AF216667.1	AL359758.1	AL589764.2		
AL589765.1	AL589986.2	AL590431.1	AL590431.2	AL590431.3	AL590666.2	AL591893.1	AL591893.2	AL591893.3	AL592307.1		
AL592492.2	AL606500.1	AL713999.1	ANTXRL	AP000304.12	AP000313.1	AP000313.2	AP000569.2	AP000569.8	APOA1BP		
AQP10	ASH1L	ATP5O	ATP8B2	BCAN	BCL9	C1orf104	C1orf189	C1orf230	C1orf43	C1orf46	
C1orf61	C1orf68	CAPN2	CCDC47	CCT3	CGN	CHD1L	CHRNA2	CKS1B	CLK2	CRABP2	
CRCT1	CRNN	DAP3	DCST1	DCST2	DDX42	DPM3	EFNA1	EFNA3	EFNA4	FAM189B	FDPS
FLAD1	FLG	FLG2	FTSJ3	GABPB2	GBA	GBAP	GPATCH4	HAPLN2	HAX1	HCN3	
HDHD1CP	HELZ	HRNR	hsa-mir-190b	hsa-mir-548d-2		hsa-mir-554	hsa-mir-555	hsa-mir-9-1	hsa-mir-92b	IL6R	
IQGAP3	ITSN1	KCNH6	KCNN3	KPRP	KRTCAP2	LCE1D	LCE1E	LCE1F	LCE2A	LCE2B	
LCE2C	LCE2D	LCE3A	LCE3B	LCE3C		LCE3D	LCE3E	LCE4A	LCE5A	LLEP1	
LENEP	LIMD2	LINGO4	LOR	LRRC6	LYSMD1	MAP3K3	MEF2D	MLLT11	MRPL9	MSTO1	MTX1
MTX1P1	MUC1	NBPF10	NBPF15	NBPF16	NES	NOTCH2NL	NUP210L		OAZ3	OR13Z1P	

PBXIP1	PGLYRP3	PGLYRP4	PI4KB	PIP5K1A	PITPNC1	PKLR	PMVK	POGZ	PRR9	
PSMB4	PSMC5	PSMD12	PSMD4	PTPRD	PYGO2	RAG1AP1	RFX5	RHBG	RP11-107M16.2	RP11-126K1.2
RP11-126K1.6		RP11-126K1.8			RP11-139I14.2	RP11-201K10.1		RP11-21N7.2		RP11-21N7.6
RP11-21N7.7		RP11-243J18.2			RP11-263K19.4		RP11-263K19.6		RP11-274N19.2	
RP11-284F21.5		RP11-284F21.7			RP11-289I10.1		RP11-289I10.2		RP11-292F22.2	
RP11-292F22.3		RP11-29H23.1			RP11-29H23.2		RP11-29H23.4		RP11-29H23.5	
RP11-29H23.6		RP11-307C12.11			RP11-316M1.11		RP11-337C18.4		RP11-350G8.3	
RP11-350G8.4		RP11-350G8.5			RP11-350G8.7		RP11-404E16.1		RP1-140J1.1	RP1-140J1.4
RP11-441L11.1	RP1-14N1.2			RP11-540D14.6		RP11-540D14.8		RP11-61L14.2		RP11-61L14.6
RP11-666A1.1		RP11-666A1.2			RP11-666A1.3		RP11-666A1.4		RP11-666A1.5	
RP11-66D17.3		RP11-66D17.5			RP11-68I18.2		RP11-74C1.2	RP11-74C1.4		RP11-763B22.2
RP11-763B22.3		RP11-763B22.4			RP11-89F3.2	RP11-89F3.3	RP11-98D18.1		RP11-98D18.2	
RP11-98D18.3	RP11-98D18.7			RP11-98D18.9		RP11-98G7.1		RP1-43O17.1		RP1-43O17.2
RP1-52J10.9	RP1-91G5.1	RP1-91G5.3	RPTN	RUSC1	SCAMP3	SCNM1	SEC22B	SELENBP1	SEMA6C	
SHC1	SHE	SMARCD2	SNORA44	SNORA58	SNORA8	SNORD59	snoU13	SNX27	STRADA	
TACO1	TCAM1	TCHH	TCHHL1	TDRD10	TDRKH	THBS3	TMEM71	TMOD4	TNFAIP8L2	
TNRC4	TPM3	TRIM46	TTC24	TUFT1	U2	U6	U7	UBAP2L	UBE2Q1	VPS72
Y_RNA	YY1AP1	ZBTB7B	ZNF687							

PANTHER analysis: 137 mapped ids are found, 118 mapped ids are not found.

5.8.6.3 Genomic Gains in Oestrogen Receptor Positive DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

There were gains present in both ER positive DCIS (n=5/17) and ER positive invasive breast disease (n=7/9) with considerable overlap in this series.

1. p-values < 0.05;
2. A gene list is mapped from 3 genomic regions found on chromosomes 6, 16;
3. These regions encompass 16 genes altered in both ER positive DCIS and ER positive invasive breast disease;

Genes:

AC106739.1	AL031963.1	AL031963.2	AL031963.5	BPHL	HLA-DQA1	HLA-DQB1	IL4R	JMJD5	NSMCE1	RP1-
40E16.11	RP1-40E16.8		RP1-40E16.9							

PANTHER analysis: 10 mapped ids are found, 6 mapped ids are not found.

5.8.6.4 Genomic Sc Gains in Oestrogen Receptor Positive DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

There are similar Sc gains in both ER positive DCIS (n=4/17) and ER positive invasive breast disease (n=6/9) in this series.

1. p-values < 0.05;
2. A gene list is mapped from 16 genomic regions found on chromosomes 6, 12, 16, 18, 20;
3. These regions encompass 63 genes altered in both ER positive invasive breast disease and ER positive DCIS;

Genes:

ASXL1	5S_rRNA	ABCC11	AC106739.1	AC125494.2	AL031963.1	AL031963.2	AL031963.3	AL031963.5	AL121935.1
AL133343.1	ATN1	BPHL	C12orf57	C20orf112	CD4	COPS7A	EMG1	ENO2	FAM136B
GNB3	GPR162	HLA-DQA2	HLA-DQB2	HLA-DQB3	hsa-mir-141	hsa-mir-200c	IL4R	JMJD5	LAG3 LEPREL2
LONP2	LRRC23	MACROD2	MLF2	NQO2	NSMCE1	PHB2	PTMS	PTPN6	RIPK1
RP11-356I2.4		RP11-410N8.1		RP11-410N8.3	RP11-410N8.4		RP11-568A7.1		RP11-356I2.2
40E16.11	RP1-40E16.2		RP1-40E16.3		RP1-40E16.4		RP1-40E16.8		RP1-40E16.9
RP1-90J20.8	RP5-1184F4.5		snoU89	TCP10L2	TNFAIP3	TOP1	TUBB2A	U47924.1	U7

PANTHER analysis: 30 mapped ids are found, 33 mapped ids are not found.

5.8.6.5 Losses in Oestrogen Receptor Positive DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

There are similar losses in both ER positive DCIS (n=7/17) and ER positive invasive breast disease (n=8/9) in this series.

1. p-values < 0.05;
2. A gene list is mapped from 29 genomic regions found on chromosomes 1, 8, 13, 15, 16, 22, X;
3. These regions encompass 114 genes altered in both ER positive invasive breast disease and ER positive DCIS;

Genes:

5S_rRNA	AC022716.1	AC039056.1	AP000351.4	AP000553.1	AP000555.1	ARHGEF10	ATXN8OS	CCDC116	CLN8
CSMD1	CT45A1	CT45A2	CT45A3	CT45A4	DDT	DLC1	EHMT1	EIF4A1P6	F8A3
GANC	GFRA2		GSTT1	GSTT2	GSTTP1	H2AFB3	hsa-mir-1184	hsa-mir-130b	hsa-mir-301b
mir-596	hsa-mir-627	JPH3	KB-1027C11.4		KLHL1	LARGE	MAPK1	MRRFP	PAK3
PLA2G4D	PLA2G4E	PLA2G4F	PPIL2	PRKCZ		RP11-110K18.2		RP11-206H15.2	RP11-218L14.4
RP11-264J4.5		RP11-264J4.6		RP11-520F24.1		RP11-520F24.2		RP11-520F24.3	RP11-521H3.1
RP11-77P3.1		RP11-954J6.3		RP13-36C9.4		RP13-36C9.5		RP5-1168A5.1	RP5-
1170D6.1	RP5-964N17.1		SDF2L1	TDRD3	TMEM87A	TMLHE		TOP3B	U3
									U6
UBE2L3	VPS39	Y_RNA	YDJC	YPEL1	ZMYM2				

PANTHER analysis: 35 mapped ids are found, 37 mapped ids are not found.

5.8.6.6 Total Loss in Oestrogen Receptor Positive DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

There were total losses present in ER positive DCIS (n=13/17) not observed in ER positive invasive breast disease (n=0/9).

1. p-values < 0.05;
2. A gene list is mapped from 29 genomic regions found on chromosomes 3, 4, 8, 9, 10, 11, 13, 15, 16;
3. These regions encompass 60 genes altered in ER positive DCIS;

Genes:

7SK	AC025678.1	AC090947.1	AC108056.1	AC133041.1	AL596092.1	AP000720.1	AP001482.1	AP001970.1	ATP8A2	CEP164
	CTD-2026G6.1		CTD-2026G6.2		DSCAML1	ERN2	GRM5	hsa-mir-1324LHFP	MTUS2	
	NEFL	NEFM	NOX4	NUPL1	OR10D1P	OR10D3P	OR10V1	OR8B1P	OR8B2	OR8B3
	OR8C1P	OR8D1	OR8D2	OR8F1P	OR8G2P	OR8G5OSBP	PATL1	PLK1	PPARG	RP11-206I15.2
	RP11-241K7.1	RP11-241K7.2		RP11-351N4.3		RP11-413E6.1		RP11-413E6.2		RP11-413E6.3
	RP11-468L18.1		RP11-527F15.1		RP11-642N14.3		RP11-803B1.1		RP11-803B1.2	
	RP11-803B1.3		RP11-803B1.4		SNORA31	TYR	U6	U7	U8	VWA5A

PANTHER analysis: 22 mapped ids are found, 38 mapped ids are not found.

There were total losses present in ER positive invasive breast disease (n=9/9) not observed in ER positive DCIS (n=0/17).

1. p-values < 0.05;
2. A gene list is mapped from 23 genomic regions found on chromosomes 9, 13, 14, 16, X;
3. These regions encompass 120 genes altered in ER positive invasive breast disease;

Genes:

5S_rRNA	AC002504.1	AC112778.1	AC112778.2	AL136001.1	AL136359.1	AL138690.1	AL139082.1	AL162377.1	
	AL162377.2	AL163544.1	AL355611.1	AL355611.2	AL355611.3	AL356863.1	AL359180.1	AL391384.1	AL442203.2
	AL512362.1	AL590007.2	ARX	ASPG	ATP7B	ATXN8OS	C13orf23	C13orf34	C13orf36
	C14orf180	CCDC70	CCNA1	DACH1	DGKH	DIS3	DUSP21	FOXO1	FREM2
	FREQ	HMCN2	hsa-mir-203	hsa-mir-759	KDM6A	KIAA0564	KIF26A	KLF12	KLF5
	KLHL1	NEK3	NEK5	NHLRC3	OLFM4	PCDH8	PIBF1	POLA1	RP11-117N4.1
	RP11-11C5.1	RP11-12I24.2	RP11-12I24.3	RP11-133M24.1	RP11-16D22.1	RP11-187A9.2	RP11-187A9.3	RP11-196I2.1	RP11-206H15.2
	RP11-206L1.2	RP11-209P2.1	RP11-214O11.1	RP11-214O11.2	RP11-218I21.1	RP11-248G5.3	RP11-24H2.1	RP11-24H2.2	RP11-266E6.2
	RP11-279N8.1	RP11-30C8.1	RP11-30C8.2						

RP11-327P2.3	RP11-335N6.1	RP11-349O10.1	RP11-381L18.2	RP11-393H6.2	RP11-393H6.3
RP11-430K22.1	RP11-430K22.2	RP11-431O22.2	RP11-459J23.1	RP11-459J23.2	
RP11-474L7.4	RP11-501G6.1	RP11-505F3.2	RP11-505F3.4	RP11-520F24.1	RP11-520F24.2
RP11-520F24.3	RP11-56M2.1	RP11-571G1.1	RP11-571G1.2	RP11-629O4.1	
RP11-76N11.2	RP11-77P3.1	RP11-7B3.2	RP11-7B3.3	RP11-7B3.4	RP11-88G17.6
SNORA25	SNORA68	SNORA9	SNORD37	STOML3	TDRD9
TMEM179	TRPC4	U4	U6	U7	
UFM1	UTP14C	Y_RNA			

PANTHER analysis: 31 mapped ids are found, 89 mapped ids are not found.

5.8.6.7 CdLOH in Oestrogen Receptor Positive DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

There is similar CdLOH present in both ER positive DCIS (n=10/17) and ER positive invasive breast disease (n=7/9).

1. p-values < 0.05;
2. A gene list is mapped from 67 genomic regions found on chromosomes 1, 2, 5, 7, 8, 9, 11, 13, 15, 19, 22, X;
3. These regions encompass 215 genes altered in both ER positive DCIS and ER positive invasive breast disease;

Genes:

5S_rRNA	7SK	AC000035.3	AC000041.8	AC002472.8	AC002472.9	AC004019.10	AC004019.16	AC004019.17
AC005300.5	AC006946.12	AC006946.15	AC008162.1	AC011890.1	AC019176.2	AC022613.1	AC022716.1	AC027807.1
AC039056.1	AC068286.1	AC074389.1	AC074389.6	AC074389.7	AC074389.9	AC090510.1	AC102941.1	AC103965.1
AC110781.5	ACRV1	ADAMTSL3	ADRBK2	AIFM3	AL022323.1	AL121825.1	AL451005.1	AP000347.2
AP000350.6	AP000350.7	AP000350.8	AP000351.3	AP000351.4	AP000553.1	AP000555.1	AP1B1	AP3B1
ATP1B4	BPIL2	C1GALT1C1	C22orf28	C22orf30	CCDC116	CDAN1	CECR1	CECR2
CECR3	CECR4	CECR5	CECR6	CFP	CHEK1	CPXCR1	CSMD1	CT45A1
CT45A2	CT45A3	CT45A4	CT47B1	CTA-211A9.5	CTA-221G9.11	CTA-256D12.11	CTA-256D12.12	CTA-292E10.6
CTA-292E10.7	CTA-407F11.6	CTA-407F11.7	CTA-407F11.8	CTA-85E5.7	CTA-984G1.5	CTD-2037K23.1	CTD-2037K23.2	
CTD-2254N19.1	CUL4B	CXorf25	CXorf64	DDT	DDTL	DLC1	EFCAB6	EHMT1
EIF4A1P6	ELFN1	ELK1	EMID1	FAM189A1	FAM32B	FAM70A	FBXO7	GANC
GFRA2	GRIA3	GSTT1	GSTT2	GSTT2B	GSTTP1	HAUS2	HMG2L9	HORMAD2
hsa-mir-130b	hsa-mir-301b	hsa-mir-627	IGLV10-67	IGLV1-62	IGLV1-63	IGLV1-64	IGLV1-65	IGLV1-66
IL17RA	KB-1027C11.4	KIAA1671	LAMP2	LARGE	LL22NC03-23C6.15	LL22NC03-28H9.5	LRRC57	LZTR1
MAD1L1	MAPK1	MCTS1	MRRFP	MTMR3	NEFH	NF2	NIPSNAP1	PA2G4P1
PAK3	PATE1	PATE2	PDE4D	PLA2G4D	PLA2G4E	PLA2G4F	PNPLA5	PPIL2
PRAMEL	PRKCZ	RFPL1	RFPL1S	RFPL3	RFPL3S	RGL4	RHBD3	RP11-137D19.1
RP11-442I12.1	RP11-45J1.1	RP1-149A16.12	RP1-149A16.15	RP1-149A16.16	RP1-149A16.17	RP1-149A16.3	RP11-521H3.1	RP1-212G6.4
RP1-222H5.1	RP1-302D9.2	RP1-302D9.3	RP1-302D9.4	RP1-302D9.5	RP1-30E17.1	RP1-30E17.2		

RP13-36C9.4	RP13-36C9.5		RP3-388M5.8		RP3-393P12.1		RP4-655L22.2		RP4-655L22.4
RP5-1168A5.1		RP5-1170D6.1		RP5-964N17.1		SC22CB-1D7.1	SCAMP1		SDF2L1
SGSM1	SLC27A6	SNAP23	SNORA70	SNORD56	snoU13	STARD9	SULT4A1	SYN1	SYN3
TDRD3	THAP7	THOC5	TJP1	TMEM211	TMEM87A	TOP3B	TTBK2	TTC28	U3
									U6
UBE2L3	UBR1	UXT	VPS39	XXbac-B135H6.15		Y_RNA	YDJC	YPEL1	Z71183.1
									ZBTB33

ZNRF3

PANTHER analysis: 96 mapped ids are found, 115 mapped ids are not found.

5.8.6.8 CnLOH in Oestrogen Receptor Positive DCIS Compared to Oestrogen Receptor Positive Invasive Breast Disease

There is CnLOH in ER positive DCIS (n 16/17) not observed in ER positive invasive breast disease (n=0/9).

1. p-values < 0.05;
2. A gene list is mapped from 36 genomic regions found on chromosomes 1, 3, 4, 12, 19, 20;
3. These regions encompass 85 genes altered in ER positive;

Genes:

Mar-02	5S_rRNA	7SK		AC003111.1	AC003111.2	AC006128.1	AC008764.1	AC022506.2	AC024075.1	AC027667.1	AC027667.2
	AC034193.1	AC099680.1	AGAP2	AK3L1	AKAP8	AKAP8L	AL121988.1	BRD4	C19orf42	C19orf44	
C1orf212		C3orf10	CALR3	CHERP	CST1	CST2	CST4	CST5	CSTP1	CSTP2	CTD-2356P16.1
	CYP2A13	CYP2F1	CYP2S1	DLGAP3	ELAVL4	EPHX3	GJA4	GJB3	GJB4	GJB5	GSTM3P
	HNRNPM	hsa-mir-1470	IRAK2	KIF5A	MACROD2	MED26	NKX2-2	NKX2-4		NOTCH3	OS9
	PIP4K2C	PLK1S1	PRAM1	RASAL3	RP11-227D2.3	RP11-415I12.1		RP11-492I2.1			RP11-766N7.1
	RP1-198K11.3		RP4-777D9.2		RP5-1177M21.1		RP5-872K7.2		RP5-872K7.7		RP5-
984P4.1	RP5-984P4.4		RP5-997D16.2		RPL24P2	RPS15AP1	SLC35E1	snoU13	TBC1D30	TMEM38A	U1
	U6	U6atac	VHL	WIBG	WIF1	WIZ	XRN2	XXyac-YX60D10.1			

PANTHER analysis: 42 mapped ids are found, 43 mapped ids are not found.

There is CnLOH present in ER positive invasive breast disease (n=9/9) not observed in ER positive DCIS (n=0/17).

1. p-values < 0.05;
2. A gene list is mapped from 16 genomic regions found on chromosomes 1, 3, 11, 12;
3. These regions encompass 60 genes altered in ER positive DCISv

Genes:

5S_rRNA	AC012598.1	AC044839.1	AC044839.2	AC061999.1	AC092910.1	AC092910.2	AC103681.1	AC103681.2	AC117377.1
	AC119795.1	ADPRHL2	ALX4	ANKS1B	C11orf74	C11orf94	C3orf15	COL8A2	COX17
	DDIT3	EIF2C3	EXT2	GAP43	GSK3B	GYLTL1B	hsa-mir-616	KIF5A	LSAMP
	MARS	MBD6	MGAT4C	MIP NR112	PEX16	PHF21A	POPDC2	PRDM11	RAG1
	RP11-169N13.1	RP11-169N13.4		RP11-18H7.1		RP11-190C22.1		RP11-326J18.1	RP11-767L7.1
	RP11-767L7.2	RP4-665N4.4		RP4-665N4.5		SLC35C1	SYT13	TEKT2	TIMELESS
	TRAF6	TSPAN18	U6						TP53I11

PANTHER analysis: 37 mapped ids are found, 23 mapped ids are not found

5.8.7 Copy Number Aberrations for Triple Negative DCIS Compared to Non Triple negative DCIS

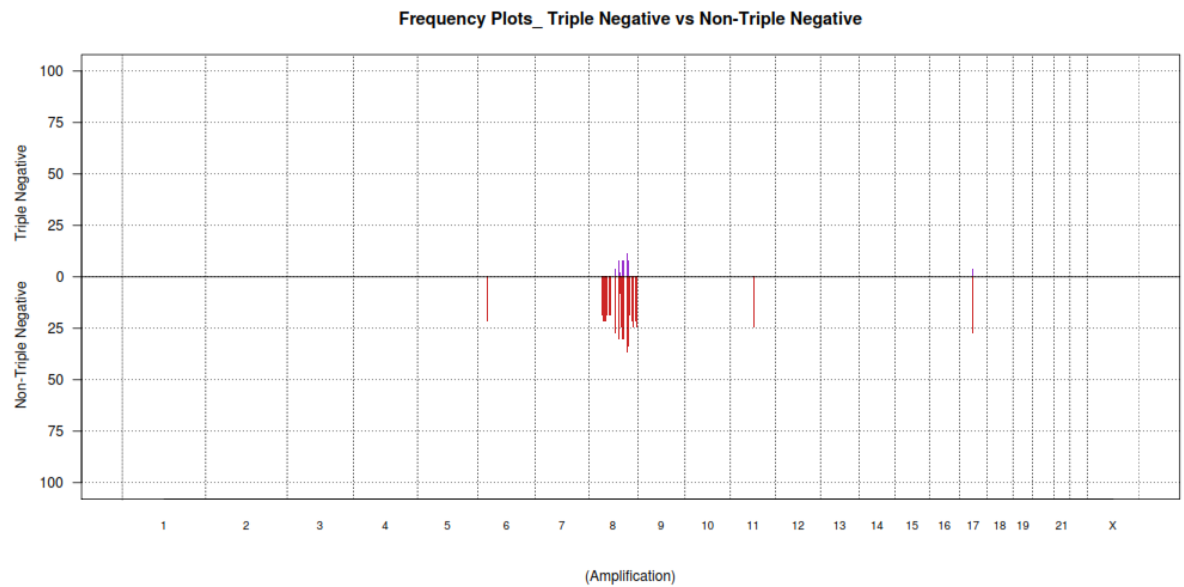
This series examines the differences between pure triple negative DCIS and all other types of DCIS (e.g. ER+, HER2+) (pure and in the presence of invasive disease).

Frequency plots showing copy number aberrations between TN DCIS and non TN DCIS are given in Figure 40.

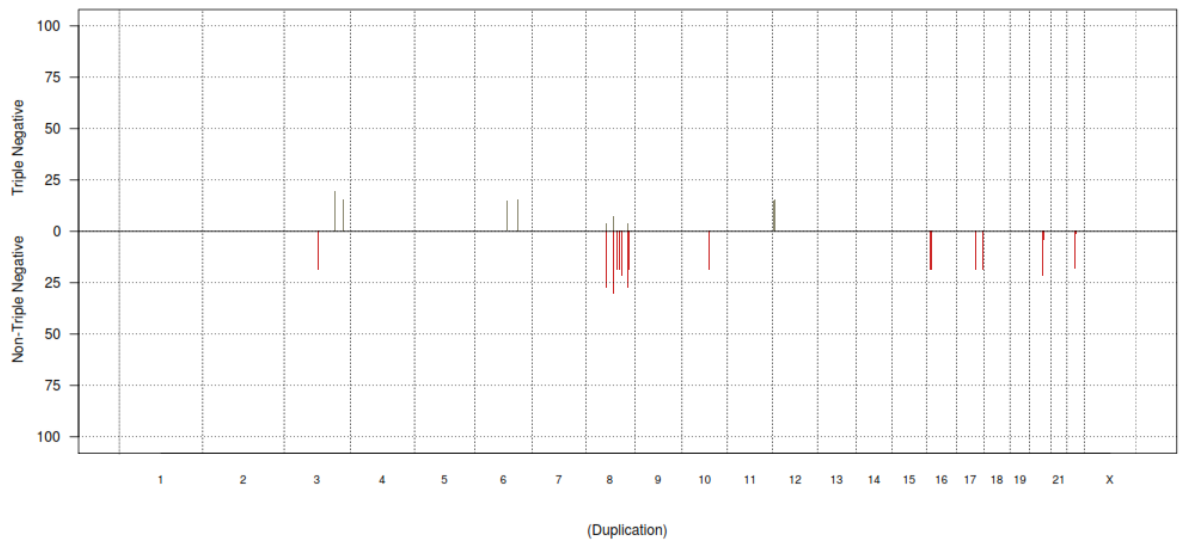
Figure 40: Frequency plots for copy number aberrations for triple negative DCIS versus non triple negative DCIS (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH, pages 220-223).

Frequency plots below show the relative copy number gains and losses for an entire sample set compared to its counterpart e.g. the first plot in Figure 36 shows amplification for all triple negative DCIS and invasive breast disease compared to all non triple negative DCIS and invasive breast disease in this series. The number on the y axis refers to the number of cases showing a genomic difference. The x axis refers to the chromosome number. The lower half of the y axis represents the non triple negative and the upper half the triple negative. Amplification on chromosome 8 is apparent on several regions for non triple negative DCIS and invasive breast disease and is also present at a lower frequency on fewer loci for triple negative DCIS

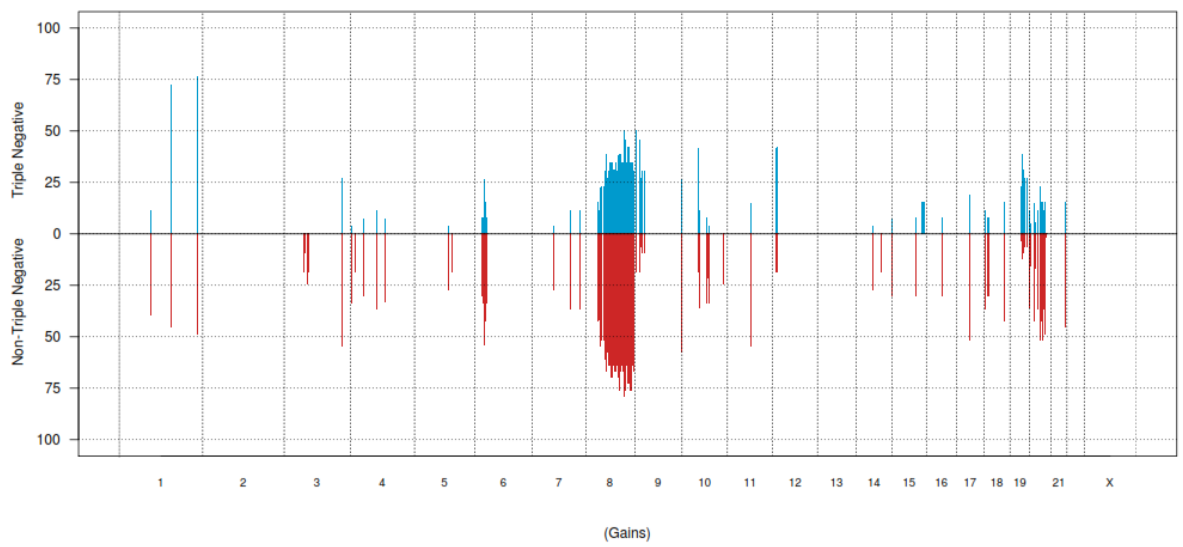
and invasive breast disease. Amplification is present on chromosomes 6 (n=24), 11(n=25) and 17 (n=26) for non triple negative DCIS and invasive breast disease but absent in triple negative DCIS and invasive breast disease.



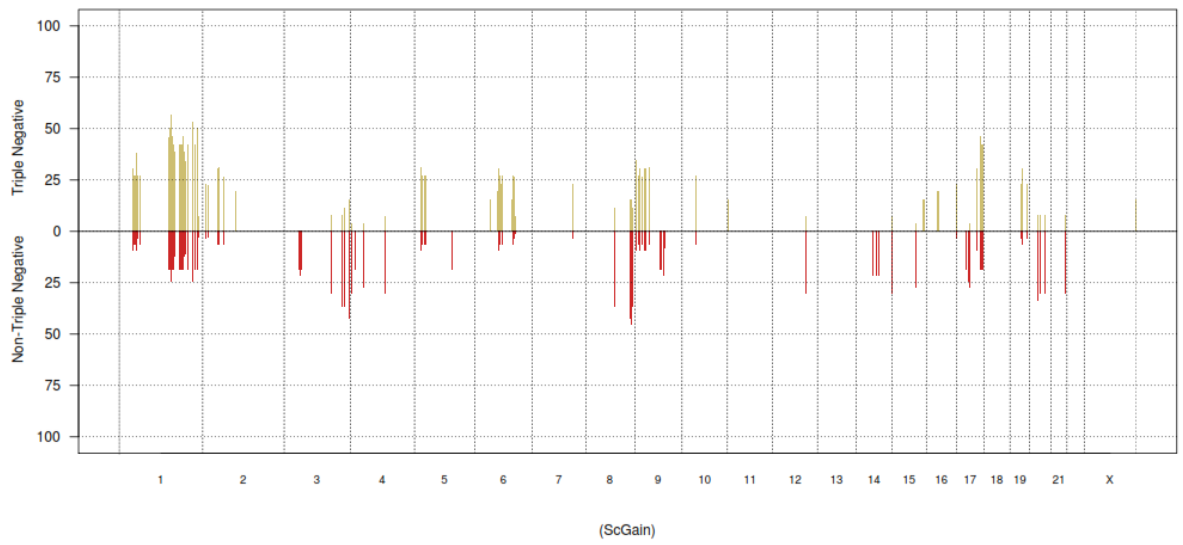
Frequency Plots_ Triple Negative vs Non-Triple Negative



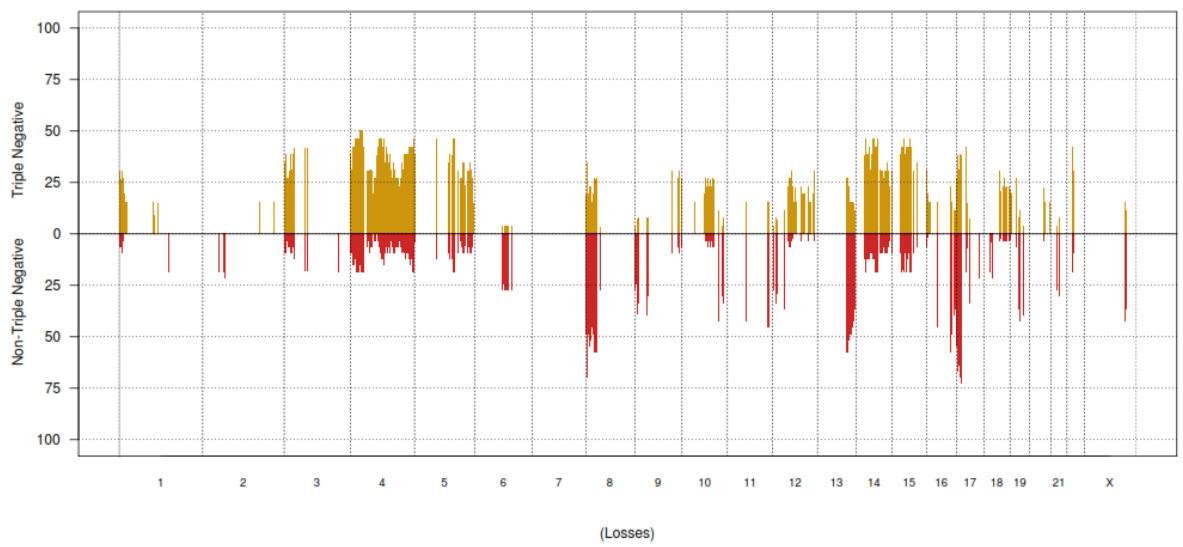
Frequency Plots_ Triple Negative vs Non-Triple Negative

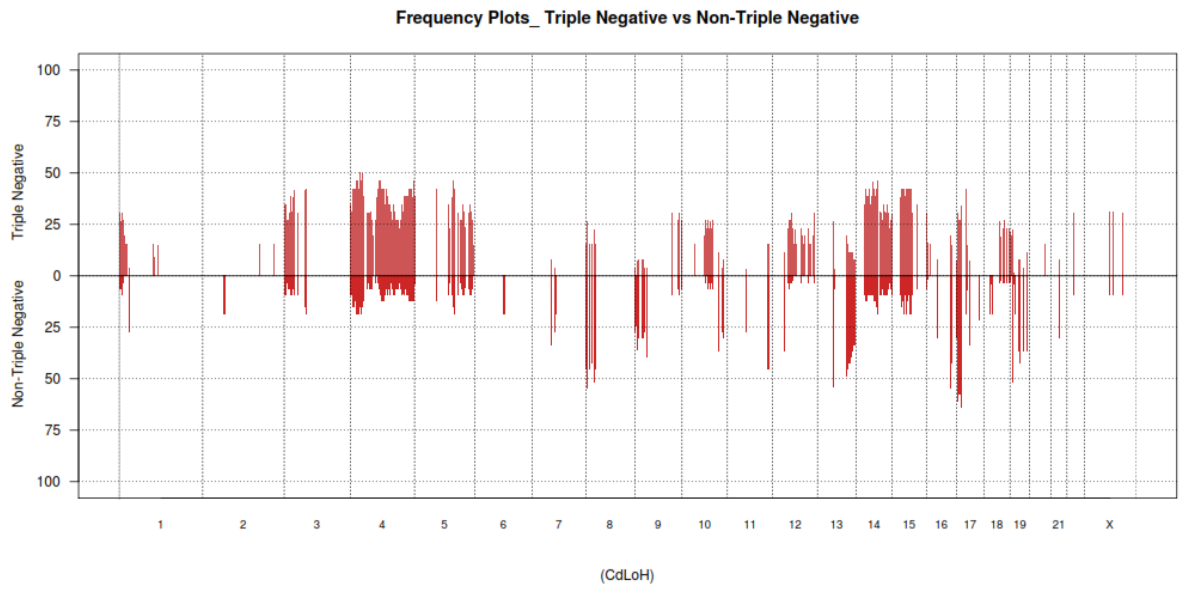
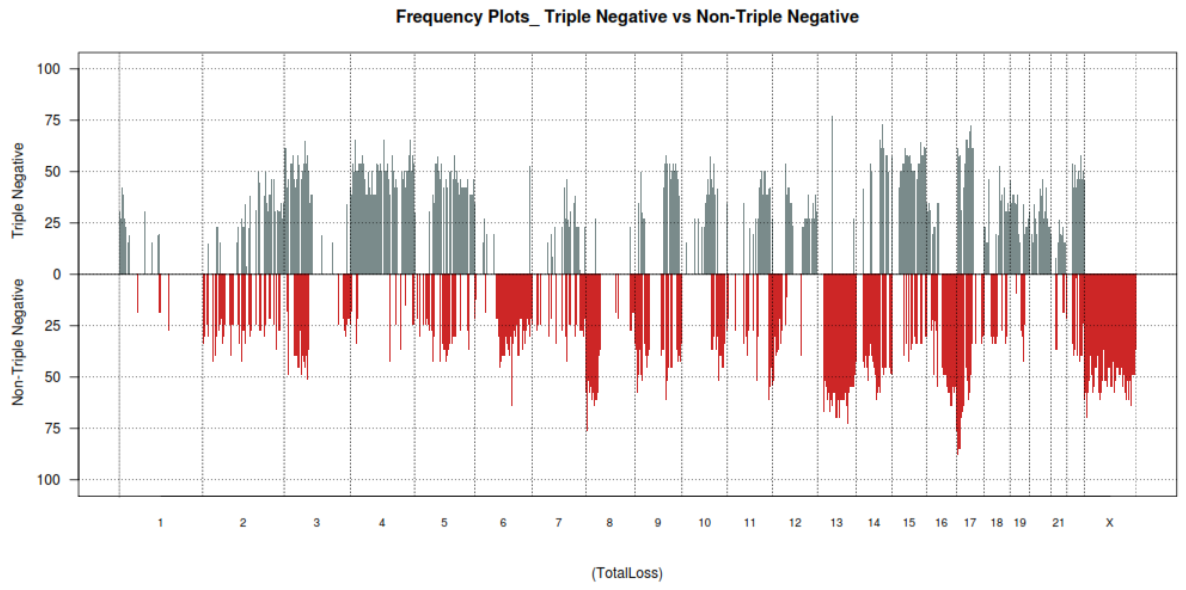


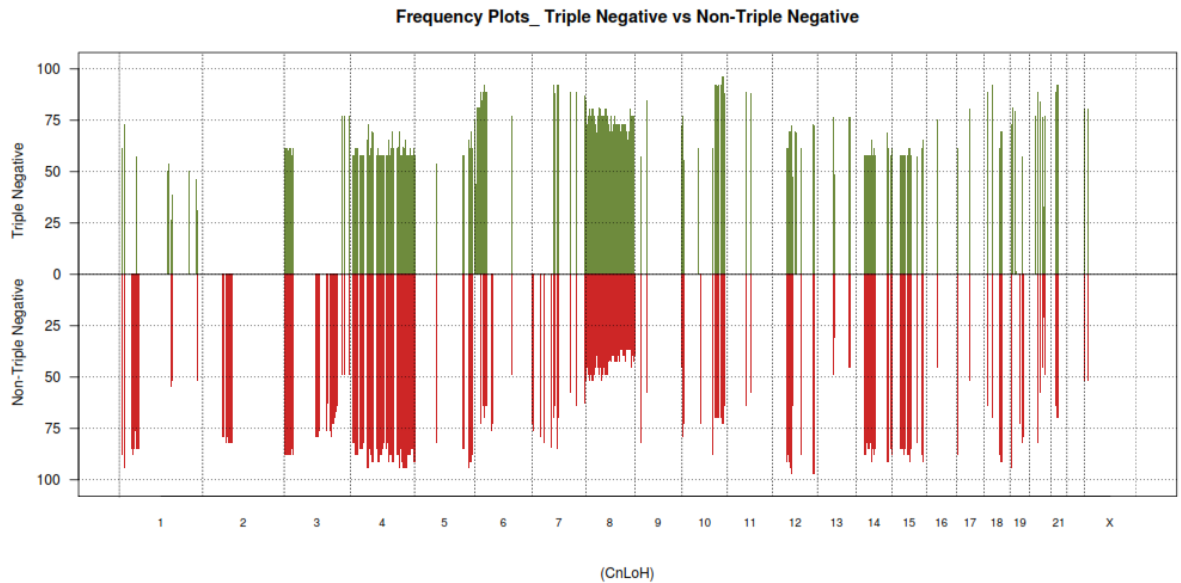
Frequency Plots_ Triple Negative vs Non-Triple Negative



Frequency Plots_ Triple Negative vs Non-Triple Negative







PANTHER Analysis is used to provide the following information the above copy number scenarios for all groups listed in sections 5.3.3 to 5.3.8

1. p-values- a significant p-value showing the genomic aberrations are not defined by chance and that the null hypothesis is correct, as an example – duplications found in triple negative DCIS compared to triple negative DCIS- a p-value of 0.5 showing that the there is a 5% chance that these duplication differences and found by random chance alone within our series.
2. The gene list mapped from different genomic regions on each chromosome. Each chromosome has been divided into overlapping genomic regions which are assessed for amplifications, duplications etc. Genes are identified in each region from different chromosomes. For each analysis the number of gene regions and chromosomes is given.

3. A list of genes is created from an online database (ESEMBL) and these genes are listed in the results section as gene ids.
4. The gene list (ids) are uploaded to the PANTHER analysis software, homo sapiens is the species genome selected and a functional analysis is requested. The Software provides a list of mapped genes with known biological, molecular and protein functions. Those genes entered into the software with established functions are returned as “mapped ids found” and those without are returned as “mapped ids not found”. The mapped ids are used to create pie charts showing biological processes, molecular functions and radar graphs for protein classes (representative charts and graphs are found in 5.3.7 the majority are stored on the accompanying disc.

Genomic data and subsequent PANTHER analysis for TN negative DCIS and non triple negative DCIS is given below.

5.8.7.1 Amplification in Triple Negative DCIS Compared to Amplification in Non Triple Negative DCIS

The comparison of amplifications in TN DCIS (n=3/26) compared to non TN DCIS (n=12/33) shows considerable overlap between those amplifications found in TN DCIS also found in non TN DCIS.

1. p-values < 0.05;
2. A gene list is mapped from 19 genomic regions found on chromosomes 6, 8, 11, 17;
3. These regions encompass 39 genes altered in TN DCIS and non TN DCIS;

Genes:

AC107374.1 ADCY8 AF216667.1 CCND1 CEBPD CRKRS CSMD3 DENN DENND3 HIST1H2AB HIST1H3A HIST1H3B HIST1H4A
HIST1H4B hsa-mir-151 IKBKB KIAA0146 KIAA1429 LRRC6 MYEOV ORAOV1 POLB PRKDC PTK2 PTK2 PTK2 PXDNL RP11-231D20.1
RP11697N18.1 SLC45A4 SNORA7 U6 U6

PANTHER analysis: 19 mapped ids found, 10 mapped ids are not found.

Amplifications are found in non TN DCIS (n=8/33) that are not observed in TN DCIS (n=0/26)

1. p-values < 0.05;
2. A gene list is mapped from 18 genomic regions found on chromosomes 8, 11, 17;
3. These regions encompass 39 genes altered in non TN DCIS;

Genes:

7SK AC013300.1 AC022973.1 AC103833.1 AC109322.1 ADCY8 C8orf4 CRKRS CSMD3 CYC1 DEPD6 EFR3A EXOSC4
FAM49B FGF19 FGF3 FGF4 GPAA1 GPR20 GRINA GSDMC KCNQ3 KIAA1875 LRRC6 MAF1 OC90 OPLAH PARP10 PLEC PTK2 RP11-
1023P17.1 RP11274M4.1 RP11473O4.1 SHARPIN SLC45A4 SNORA25 SPATC1 ST18 TMEM71

PANTHER analysis: 27 genes mapped, 1 id is not found.

5.4.1.1 *Duplication of Triple Negative DCIS Compared to Duplication in Non Triple Negative DCIS*

Duplications are found in TN DCIS (n=5/26) that are not observed in non TN DCIS (n=0/33)

1. p-values < 0.05;
2. A gene list is mapped from 18 genomic regions found on chromosomes 3, 6, 12;
3. These regions encompass 17 genes altered in TN DCIS;

Genes:

5S_rRNA AC004671.1 AC106722.1 AL096711.1 C6orf174 ECHDC1 MBNL1 NTF3 RNF146 RP11362A9.3 RP1-177A13.1 RP3-
351K20.3 RP3-351K20.4 RP3-403A15.1 RSPO3 TMEM14E Y_RNA

PANTHER analysis: 6 mapped ids found, 11 mapped ids are not found.

Duplications are found in non TN DCIS (n=0/33) that are not observed in TN DCIS (n=7/26).

1. p-values < 0.05;

2. A gene list is mapped from 20 genomic regions found on chromosomes 3, 8, 10, 16, 17, 20, 22;
3. These regions encompass 54 genes altered in non TN DCIS;

Genes: AC015813.2AC020688.1 AC090922.1 AC092291.1 AC124319.3 AC136443.1 AC138932.2 AL031661.1 AL391421.1
 AL391421.2 AL391421.3 AL391421.4 AP003355.2 ATG3P1 BFAR CTD-2561B21.1 CUEDC1 DLG5 FAM18A
 GSTTP2 H2AFZP5 hsamir-1972 HSPEP1 KCNS2 MAFB MRPS23 NIT2 NOMO1 NPIP NUBP1 OXR1 PARN PDXDC1 PLA2G10
 POLR3A RALGAPB RP11-101E14.2 RP11-101E14.3 RP11126H7.3 RP1-191L6.2 RP4-564F22.2 RP4 564F22.5
 RP5-1031J8.1 RPS24 SNHG11 SNORA71 SNORA71A SNORA71B SNORD112 TBC1D23 U6 U7 VEZF1
 Y_RNA

PANTHER analysis: 19 mapped ids are found, 35 mapped ids are not found.

5.8.7.2 Genomic Gains in Triple Negative DCIS Compared to Genomic Gains in Non Triple Negative DCIS

Gains found in TN DCIS (n=10/26) have some overlap with gains found in non TN DCIS (n=3/33).

1. p-values < 0.05;
2. A gene list is mapped from 17 genomic regions found on chromosomes 3, 4, 5, 10, 10;
3. These regions encompass 77 genes altered in non TN DCIS;

Genes:

5S_rRNA AC006296.3 AC022400.3 AC096971.2 AC103560.1 ACADSB AGAP5APC BMS1P4 BUB3 C10orf55 CAMK2G CHCHD1
 COQ6CTC-459M5.1 CTC-459M5.2 CTC-493D22.1 CTD2077I5.1 CTD 2192A1.1CTD-2201G3.1 ENTPD5 EPB41L4A FAM161B FHIT
 FOXP1 FTHL23 FUT11 GLT8D4 GLUDP3 HMX2 HMX3 IKZF5 KIAA0913 MCC MYOZ1 NCRNA00219 NDST2 PLAU
 PPP4R2 PSTK PTGR2 RNase_MRP RP11-162A23.5 RP11-282I1.1 RP11-417O11.5 RP11430J3.1 RP11-430J3.2
 RP11-458B24.2 RP11-464F9.1 RP11-464F9.19 RP11-464F9.9 RP11526F3.1 RP11-574K11.12 RP11-574K11.16 RP11-574K11.18
 RP11-574K11.20 RP11-574K11.5 RP11-574K11.8 RP11-624D20.1 RP11-654C22.1 RP11-728B21.2 RP11-905F6.1 SEC24C SHQ1
 SNORA13 snoU13 SYNPO2L TSSK1B U1 U4atac U6 U7 USP54 VCL Y_RNA YTHDC2 ZNF410

There were gains found in non TN DCIS (n=8/33) that are not observed in TN DCIS (n=0/26).

1. p-values < 0.05;

2. A gene list is mapped from 36 genomic regions found on chromosomes 2, 6, 7, 9, 10, 17, 18, 19;
3. These regions encompass 99 genes altered in non TN DCIS;

Genes:

AC004603.4 AC005176.1 AC005205.2 AC005205.3 AC005205.4 AC005239.1 AC006133.3 AC007842.1 AC008555.5 AC008655.1
AC009711.1 AC011443.1 AC011445.1 AC011455.1 AC011496.1 AC011500.1 AC011500.2 AC012309.4 AC012309.5 AC080032.1
AC091078.1 AC091078.2 AC092296.1 AC092296.3 AC093063.1 AC103996.1 ACAN ASPDH C19orf63 CDKN2BAS
CEACAMP7 CLC DENND4C DLL3 DYRK1B EID2 EID2B FAM71E1 FBL FBXO27 FCGBP GMFG IL28A
IL28B IL29 JOSD2 KCNC3 KCTD15 LEUTX LGALS13 LGALS14 LRFN1 LRRC4B MED29 MLLT3
MRPS12 MTAP MYBPC2 MYH14 NAPS A NAPS B NCCRP1 NFKBIB NR1H2 PAF1 PAK4 PLEKHG2
POLD1 PSG1 PSG10 PSG11 PSG6 PSG7 PSMC4 RINL RP11513M16.5 RPS16
RPS6 SAMD4BSARS2 SIRT2 snoU13 SPIB SUPT5H SYCN SYT3 TIMM50 VN1R96P Y_RNA ZFP14 ZFP36
ZFP82 ZNF146 ZNF30 ZNF383 ZNF546 ZNF565 ZNF780A ZNF780B

PANTHER analysis: 59 mapped ids found, 40 mapped ids are not found.

5.8.7.3 Genomic Sc Gains in Triple Negative DCIS Compared to Genomic Sc Gains in Non Triple Negative DCIS

There are Sc Gains found in TN DCIS (n=7/26) that are not observed from non TN DCIS (n=0/33).

1. p-values < 0.05;
2. A gene list is mapped from 21 genomic regions on 6 chromosomes 1, 2, 6, 9, 16, 19;
3. These regions encompass 86 genes altered in TN DCIS;

Genes:

5S_rRNA AC016699.1 AC017099.3 AC018892.1 AC018892.3 AC018892.4 AC018892.5 AC018892.6 AC018892.7 AC018892.8
AC018892.9 AC079337.1 AC079395.1 AC092683.1 AC092683.2 AC092683.3 AC106785.1 AC106785.10 AC106785.14 AC106785.18
AC106785.20 AC106785.22 AC106785.25 AC106785.27 AC106785.28 AC106785.6
AC116553.1 AC137932.1 AC137932.3 AC159540.1 AC159540.14 AC159540.2 AC159540.3 ACTR1B AL391807.1 AL445568.1
AL590785.1 ANKRD11 ANKRD36 ANKRD36B B3GAT2 BAI3 CDKN2BAS CEBPG CHST8 COX5B CTD-2329C7.1 DBIP1 ELOVL4 FAHD2B
FAM178B IRAK1BP1 KCNQ5 KCTD15 MTAP PEPD Phip RIMS1 RP11-135M8__A.1 RP11-173D14.3 RP11-321C1.1
RP11-35J1.1 RP11-361F19.1 RP11-380M3.1 RP11-406O16.1 RP11-789A7.1 RP1-20N4.1 RP1-232L24.3 RP3-357D13.1
RP3-357D13.3 RP3-390M24.1 RP3-474G15.1 RP3-474G15.2 RP3-485C17.1 RP3-485C17.2 RP3-525N10.2 SNORD112 snoU13
TMEM131 U3 U4 U6 U7 UBTF6 VWA3B ZAP70

PANTHER analysis: 21 mapped ids are found, 65 mapped ids are not found.

There are Sc Gains in non TN (n=8/33) not observed in TN DCIS (n=0/26).

1. p-values < 0.05;
2. A gene list is mapped from 36 genomic regions found on chromosomes 3, 4, 5, 9, 14, 17;
3. These regions encompass 93 genes altered in TN DCIS;

Genes:

7SK AC006296.3 AC068400.1 AC098869.1 AC099668.5 AC104996.1 AC105935.1 AC121252.2 AC126327.1 AC126327.4 AL117692.1
AL117692.2 AL136294.1 AL161626.1 AMIGO3 APC APEH ATP6V0A1 BLMH C14orf135 C14orf182 C14orf183 C3orf54 C9orf40 C9orf41
C9orf95 CAMKV CCDC55 CDCP1 CDH29 COASY CPD CTB-75G16.1 CTC-459M5.1 CTC-459M5.2 CTD-2184C24.1 CTD-2192A1.1
DDX52 DHRS11 DHRS7 DUSP14 EFCAB5 EPB41L4A FAM134C GCNT1 GGNBP2 GMPPB hsa-mir-423 HSD17B1
HSD17B1P1 IP6K1 LRRC9 MLX MRM1 MST1 MYO19 NAGLU NCRNA00219 NTRK2 OSTF1 PIGW PLEKHH3
PPM1A PRKAR2A PRUNE2 PSMC3IP RAD51L1 RBM47 RNF123 RP11-197P3.1 RP11-197P3.4 RP11-197P3.5
RP11-214N16.2 RP11-422N19.3 RP11-458B24.2 RP11-526F3.1 RP11-58E21.1 RP11-66C24.1 RP13-1056D16.2
RP13-480C15.1 SLC6A4 SNORA13 SNORD63 SYNRG TADA2A TMIGD1 TRAIP TRPM6 TUBG1 TUBG2 U6 UBA7 Y_RNA

PANTHER analysis: 50 mapped ids are found, 43 mapped ids are not found.

5.8.7.4 Losses in in Triple Negative DCIS Compared to Genomic Total Loss in Non Triple Negative DCIS

There are losses found in TN DCIS (n=11/26) that are not observed in non TN DCIS (n=0/33).

1. p-values <0.0001;
2. A gene list is mapped from 23 genomic regions on chromosome 4;
3. These regions encompass 94 altered genes in TN DCIS;

Genes:

MAR01 7SK AC079349.1 AC095046.1 AC095047.1 AC096766.1 AC097534.1 AC097534.2 AC104082.1 AC104082.2
AC105285.2 AGPAT9 ANXA3 ARHGAP24 C4orf12 C4orf39 C4orf43 CDS1 CPE FAM175A FGF5 FRAS1
GALNT7 GALNTL6 GK2 GK3P HMGB2 hsa-mir-1979 hsa-mir-578 KLHL2 NACA3P NKX6-1 NPY5R OR7E94P RP11-10K16.1
RP11112L18.1 RP11-147K21.1 RP11-218C23.1 RP11-219C20.1 RP11-234K19.1 RP11-234K19.2 RP11-274J2.1
RP11-294O2.1 RP11-294O2.2 RP11-366M4.1 RP11-366M4.11 RP11-366M4.12 RP11-366M4.13 RP11-366M4.14
RP11-366M4.15 RP11-366M4.17 RP11-366M4.2 RP11-366M4.3 RP11-366M4.4 RP11-366M4.5 RP11-366M4.6
RP11-366M4.8 RP11-42A4.1 RP11-436A7.1 RP11-443J23.1 RP11-452C8.1 RP11-469O11.1
RP11-469O11.2 RP11-485C11.1 RP11-489G11.1 RP11-502M1.1 RP11-502M1.2 RP11-51G24.1
RP11-606P2.1 RP11-610O8.1 RP11-689K5.2 RP11-689K5.3 RP11-722P15.1
RP11-75A5.1 RP11-767N15.1 RP11-798M19.3 RP11-8L2.1 RPL35AP12 SAP30 SC4MOL SCRG1
SNORA31 SNORA75 SPOCK3 TKTL2 TMEM192 TRIM60 TRIM61 TRIM75 U4 U5 U6 WDFY3 Y_RNA

PANTHER analysis: 23 mapped ids are found, 71 mapped ids are not found.

There are losses in non TN DCIS (n=10/33) that are not observed in TN DCIS (n=0/26) represents losses for non TN DCIS (N=10/24).

1. P values < 0.05;
2. A gene list is mapped from 35 genomic regions found on chromosomes 1, 2, 3, 6, 9, 10, 17, 18, 19;
3. These regions encompass 127 altered genes in TN DCIS;

Genes:

5S_rRNA AC002539.1 AC003051.1 AC004562.1 AC005176.1 AC005205.2 AC005205.3 AC005205.4 AC005208.1 AC005239.1
AC005307.1 AC005616.1 AC006133.3 AC007392.3 AC007392.4 AC007403.1 AC007403.2 AC007403.3 AC009474.1
AC009474.2 AC011443.1 AC011500.1 AC011500.2 AC023115.3 AC078941.1 AC091038.2 AC092155.1 AC092155.2
AC092155.4 AC093063.1 AL162419.2 AL356957.10AL356957.2 AL356957.3 AL356957.4 AL356957.6 AL356957.7
AL356957.9 AL359836.1 AL732363.1 C9orf68 CAPN12 CDH2 CLCDLL3 DYRK1B EHBP1 EID2
EID2B EMX2OS FAM98C FBL FBXO27 FCGBP FSHR GLIS3 GMFG hsa-mir-1-2 hsa-mir-133a-1
KCNJ16 KCNJ2 KCNK18 KIAA1598 LEUTX LGALS13 LGALS14 LGALS7 LRFN1 MAP4K1 MED29
MEIS1 MIB1 NT5E PAF1 PAK4 PDZD8 PLEKHG2 RASGRP4 RP11-14N7.1
RP11-14N7.2 RP11-268I9.1 RP11321N4.3 RP11-33E24.1 RP11-33E24.3
RP11-358M14.2RP11-403I13.1 RP11403I13.2 RP11-403I13.4 RP11-403I13.5 RP11-403I13.6
RP11-403I13.7 RP11403I13.8RP11-501J20.2 RP11-501J20.3 RP11-501J20.5 RP11-5G18.2 RP11-70J12.1
RP11-744H18.1 RP11-744H18.2 RP11-763B22.5 RP11-763B22.6 RP11-763B22.7 RP11-763B22.8
RP11-763B22.9 RP1-3J17.3 RPS16 RSL24D1P2 RYR1 SAMD4B SARS2 SLC18A2 SLC1A1
SNHG5 SNORA73 SNORA81 SNORD50 snoU13 SNX14 SPRED3 SUPT5H SYNCRIP TIMM50 U1 U6 VAX1
Y_RNA ZFP36

PANTHER analysis: 41 mapped ids are found, 86 mapped ids are not found.

5.8.7.5 Total Loss in Triple Negative DCIS Compared to Genomic Total Loss in Non Triple Negative DCIS

There are total losses found in TN DCIS (n=17/26) that are not observed in non TN DCIS (n=0/33).

1. p-values = 0.05;
2. A gene list is mapped from 58 genomic regions found on chromosomes 3, 4,14,15,17;
3. These regions encompass 223 genes altered in Non TN DCIS;

Genes:

5S_rRNA 7SK ABCD4 AC002117.1 AC003043.1 AC003043.2 AC003102.3 AC003963.2 AC004222.1 AC004596.1
AC004968.1 AC004968.2 AC006349.1 AC007056.1 AC007686.1 AC007722.1 AC008045.1 AC008105.1 AC012409.1

	AC015936.1	AC015936.3	AC016251.1	AC019131.2	AC020704.1	AC026202.1	AC026202.5	AC026882.1	AC027708.1
	AC027807.1	AC036222.1	AC055813.1	AC055873.1	AC068400.1	AC079915.1	AC084809.2	AC084809.3	AC090043.1
	AC090283.1	AC090283.3	AC091199.1	AC096750.1	AC102948.1	AC102948.2	AC103702.1	AC103965.1	AC104260.1
	AC113211.1	AC124789.1	AC134669.1	AC135724.1	AC138150.3	AC138150.4	AC183087.1	AC183087.2	ACBD4
	ACCN1	ADAM11	ADAMTSL3	ADCK1	ADH1A	ADH1B	ADH1C	ADH4	ADH5
	ADH6	AF111168.1	AF111168.3	AHSA1	ALKBH1	ARHGAP23	ARRDC4	ASB16	ATXN7L3
	C14orf148	C14orf156	C14orf174	C14orf178	C17orf104	C17orf105	C17orf46	C17orf53	C17orf65
	C17orf93	C1QL1	C4orf37	CA10	CCDC103	CCDC43	CD300LG	CDC6	CNTN4
	CNTN6	COX6B1P2	CTD-2175M1.1		CTD-2377D24.1		CTD-2377D24.2		CTD-2377D24.4
	CTD-2377D24.6		DBF4B	DCAKD	DDX52	DHX8	DUSP3	DYNLL1P1	EFTUD2
	EIF4E	ETV4	EVI2A	EVI2B	FAM149B2	FAM169B	FAM171A2	FAM177A2	FAM187A
	FMNL1	FZD2	G6PC3	GFAP	GJC1	GPATCH8	GPR179	GRN	GSTZ1
	HDAC5	HEXIM1	HEXIM2	HIGD1B	HNF1B	HOXB13	HOXB7	HOXB9	hsa-mir-1260hsa-
mir-1469	hsa-mir-196a-1		hsa-mir-2117ISM2		ITGA2B	KIAA1737	KIF18B	KRT222	KRT24
	KRT25	LIN52	LSM12	MAP3K14	MCTP2	MEOX1	METAP1	MPP2	MPP3
	MRPL45	MYO1D	NAGS	NF1	NGB	NMT1	NR2F2	NRXN3	OMG
	PCNAP1	PLCD3	POMT2	PPY	PRAC	PYY	RAB11FIP4	RGS6	RP11-1072C15.1
	RP11-15B17.1		RP11-176P14.1		RP11-187O7.1		RP11-204C23.1		RP11-299L17.1
	RP11-332E19.1		RP11-352E6.1		RP11-352E6.2		RP11-357H14.4		RP11-369I16.1
	RP11-428B4.2		RP11-522B15.1		RP11-522B15.2		RP11-571L19.2		RP11-571L19.4
	RP11-624A4.1		RP11-696N14.1		RP11-765K14.1		RP11-799A12.1		RP11-799A12.2
	RUNDC3A	SIPA1L1	SLC25A39	SLC4A1	SMARCE1	SNORA32	SNORA46	SNORD56	snoZ178
	SNW1	SOCS7	SOST	SPACA3	SPATA8	SPRED1	SPTLC2	SRCIN1	STON2
	SYNRG	TBC1D3	TMED8	TMEM101	TMEM63C	TMEM98	TMUB2	U3	U6
	UBTF	VEGFC	VIPAR	VSX2	WIPF2	Y_RNA	ZDHHC22		U7

PANTHER analysis: 115 mapped ids are found, 108 mapped ids are not found.

There are total losses found in non TN DCIS (n=27/33) that are not observed in TN DCIS (n=0/26).

1. p-values <0.0001;
2. A gene list is mapped from 18 genomic regions on 1 chromosome (17) ;
3. These regions encompass 46 genes altered in TN DCIS;

Genes:

AC000003.1 AC027763.1 AC027763.2 AIPL1 ALOX12 AMAC1L3 ASGR1 ASGR2 BCL6B C17orf49 C17orf61 C17orf74 CCDC42 CHRNB1
CLEC10A CTD-2545G14.1 DLG4 DPH1 FBXO39 FGF11 GAS7 hsa-mir-195 hsa-mir-497 MFSD6L NLGN2 OVCA2 PIK3R5 PIK3R6
PLSCR3 POLR2A RNASEK RNF222 RP11-609D21.1 SLC16A11 SLC16A13 SPDYE4 SPEM1 TEK1 TMEM102 TNFSF12 TNFSF12-TNFSF13
TNFSF13 TNK1 U6 Y_RNA ZBTB4

PANTHER analysis: 33 mapped ids are found, 13 mapped ids are not found.

5.8.7.6 CdLOH in Triple Negative DCIS Compared to with CdLOH in Non Triple Negative DCIS

There is CdLOH found in TN DCIS (n=11/26) that are not present in non TN DCIS (n=0/33).

1. p-values <0.0001;
2. A gene list is mapped from 24 genomic regions on 2 chromosomes (4 and 5);
3. These regions encompass 91 genes altered in TN DCIS;

Genes: Mar01 7SK AC023355.1 AC023355.3 AC079349.1 AC095047.1 AC096766.1 AC097534.1 AC097534.2 AC104082.1 AC104082.2 AC105285.2 AGPAT9 ANXA3 ARHGAP24 C4orf39 C4orf43 CDS1 CPE DUT FBN1 FGF5 FRAS1 GALNT7 GALNTL6 GK2 GK3P HMGB2 hsa-mir-1979 hsa-mir-578 KLHL2 NACA3P NKX6-1 NPY5R OR7E94P RP11-10K16.1 RP11-112L18.1 RP11-218C23.1 RP11-219C20.1 RP11-234K19.1 RP11-234K19.2 RP11-274J2.1 RP11294O2.1 RP11-294O2.2 RP11-366M4.1 RP11-366M4.11 RP11-366M4.12 RP11-366M4.13 RP11-366M4.14 RP11-366M4.15 RP11-366M4.17 RP11-366M4.2 RP11-366M4.3 RP11-366M4.4 RP11-366M4.5 RP11-366M4.6 RP11-366M4.8 RP11-42A4.1 RP11-436A7.1 RP11-443J23.1 RP11-469O11.1 RP11-469O11.2 RP11-485C11.1 RP11-489G11.1 RP11-502M1.1 RP11-502M1.2 RP11-51G24.1 RP11-606P2.1 RP11-689K5.2 RP11-689K5.3 RP11-75A5.1 RP11-767N15.1 RP11-798M19.3 RP11-8L2.1 RPL35AP12 SAP30 SC4MOL SCRG1 SLC12A1 SNORA31 SPOCK3 TKTL2 TMEM192 TRIM60 TRIM61 TRIM75 U4 U5 U6 WDFY3 Y_RNA

PANTHER analysis: 25 mapped ids are found, 66 mapped ids are not found.

There is CdLOH found in non TN DCIS (n=10/26) that are not observed in TN DCIS (0/26)

1. p-values < 0.05;
2. A gene list is mapped from 36 genomic regions found on chromosomes 2, 6, 7, 9, 10, 17, 18, 19;
3. These regions encompass 107 genes altered in TN DCIS;

Genes: 5S_rRNA AC002539.1 AC003051.1 AC004562.1 AC005176.1 AC005205.2 AC005205.3 AC005205.4 AC005208.1 AC005239.1 AC005307.1 AC005616.1 AC006133.3 AC007392.3 AC007392.4 AC007403.1 AC007403.2 AC009474.1 AC009474.2 AC011443.1 AC011500.1 AC011500.2 AC091038.2 AC092155.4 AC093063.1 AL133476.1 AL359836.1 ATP8B5P C6orf161 CAPN12 CDH2 CLC DLL3 DYRK1B EHB1 EID2 EID2B EMX2OS FAM98C FBL FBXO27 FCBGP GLIS3 GMFG hsa-mir-1-2 hsa-mir-133a-1 KCNJ16 KCNJ2 KCNK18 KIAA1598 LEUTX LGALS13 LGALS14 LGALS7 LRFN1 MAP4K1 MED29 MEIS1 MIB1 NT5E PAF1 PAK4 PDZD8 PLEKHG2 RASGRP4 RP11-207F8.1 RP11-268I9.1 RP11-30P6.1 RP11-30P6.2 RP11-30P6.3 RP11-30P6.4 RP11-30P6.6 RP11-321N4.3 RP11-33E24.1 RP11-33E24.3 RP11-424A16.1 RP11-501J20.2 RP11 501J20.3 RP11-501J20.5 RP11-5G18.2 RP1-161C16.1 RP11-70J12.1 RP1-263J7.1 RP1263J7.2 RP1-3J17.3 RP3-455E7.1 RYR1 SAMD4B SARS2 SLC18A2 SNHG5 SNORA73 SNORA81 SNORD50 snoU13 SNX14 SPRED3 SUPT5H SYNCRIP TEK TIMM50 U4 U6 UNC13B VAX1 Y_RNA ZFP36

PANTHER analysis: 51 mapped ids are found, 104 mapped ids are not found.

5.8.7.7 CnLOH in Non Triple Negative DCIS Compared with CnLOH in Triple Negative DCIS

There is CnLOH found in TN DCIS (23/26) that are not observed in non TN DCIS (0/33).

1. p-values < 0.05;
2. A gene list is mapped from 31 genomic regions found on chromosomes 6,10,19, 20;
3. These regions encompass 145 genes altered in TN DCIS;

Genes:

Mar02	5S_rRNA	7SK	AC010323.1	AC114273.1	AL023583.1	AL024498.1	AL031123.1	AL031123.2	AL033539.1
AL033539.2	AL136305.1	AL136307.1	AL136307.2	AL139807.1	AL158198.1	AL158817.1	AL357054.1	AL357497.1	AL662797.1
ANGPTL4	ATP5GP1	BMP6	C6orf134	C6orf136	C6orf52	CAGE1	CAP2	CCDC90A	CD320
CD83	DHX16	DSP	DTD1	EDN1	EEF1A1P34	EEF1E1	ELOVL2	F13A1	FARS2
GCM2	GCNT2	GCNT6	HDGFL1	HIVEP1	KANK3	KIAA1949	KLF6	LASS4	LPCAT3
LY86	MAK	MRPS18B	MUTED	NDUFA7	NEDD9	NOL7	NRM	NRN1	OFCC1
PAK1IP1	PIP5K1P1	PPP1R10	PRL	PTMAP1	RAB11B	RANBP9	RBM24	RIOK1	RNF182
RP11-125M16.1		RP11-127P7.2		RP11-146I2.1		RP11-184A2.2		RP11-239H6.2	
RP11-239H6.3		RP11-288G3.2		RP11-288G3.3		RP11-288G3.4		RP11-320C15.1	
RP11-339A7.1		RP11-359N11.1		RP11-359N11.2		RP11-360O19.1		RP11-360O19.4	
RP11-360O19.5		RP11-379J5.4		RP11-379J5.5		RP11-405O10.2		RP11-456H18.1	
RP11-456H18.2		RP11-464C19.2		RP11-464C19.3		RP11-556O15.1		RP11-637O19.2	
RP11-69L16.4		RP11-69L16.5		RP11-69L16.6		RP11-716O23.1		RP11-716O23.2	RP1-
182O16.1	RP1-182O16.2		RP1-23O21.1		RP1-273P12.1		RP1-290I10.2		RP1-290I10.3
RP1-290I10.4		RP1-290I10.5		RP1-290I10.6		RP1-290I10.7		RP1-290I10.8	RP1-
303A1.1	RP1-309H15.2		RP1-62D2.3	RP1-80N2.2	RP3-336K20_B.2		RP3-380B8.1		RP3-380B8.3
RP3-380B8.4		RP3-380E11.2		RP3-398D13.2		RP3-398D13.3		RP3-429O6.1	RP3-
451B15.3	RP3-470L22.1		RP3-486D24.1		RP3-510L9.1	RP4-761I2.2	RP4-761I2.4	RP5-1068E13.3	
RPL15P3	RPS28	RREB1	snoU13	SNRNP48	SSR1	SYCP2L	TFAP2A	TMEM14B	TMEM14C
TXNDC5	U1	U6atac	U7	Y_RNA					

PANTHER analysis: 51 mapped ids are found, 104 mapped ids are not found.

There is CnLOH found in non TN DCIS (29/33) that are not observed in TN DCIS (0/26).

1. p-values < 0.05;
2. A gene list is mapped from 70 genomic regions on chromosomes 1, 2, 3, 7, 10, 19, 20;

3. These regions encompass 244 genes altered in non TN DCIS;

Genes:

5S_rRNA	AC005034.2	AC006039.1	AC007040.5	AC007099.1	AC007099.2	AC007878.1	AC010733.2	AC010733.4	AC010733.7										
	AC010744.1	AC010975.1	AC011245.1	AC011455.1	AC012494.1	AC012671.1	AC012671.2	AC016747.3	AC023881.1										
	AC062029.1	AC064872.1	AC073628.1	AC078841.5	AC083864.3	AC083864.4	AC084149.1	AC084149.2	AC092660.1										
	AC093616.1	AC096540.1	AC098581.1	AC104134.2	AC104135.2	AC104135.3	AC104135.4	AC104170.1	AC108479.1										
	AC108479.2	AC110491.1	AC117444.2	AC117444.3	ACOT11	ACPL2	ADARB2	AHSA2	AKIRIN1										
	AL031289.1	AL031985.1	AL110502.1	AL161740.1	AL592166.1	AL929472.1	AL929472.3	ARL6	ATP1B3P1										
	ATP6V1B1	BAZ1B	BEST4	BMP8B	BTBD19	C1orf122	C1orf168	C1orf175	C1orf228										
	C2orf51	C2orf74	C8A	C8B	CCDC163P	CITED4	CRYBG3	CST1	CST4										
	CSTP2	CTNNA2	CTPS	DAB1	DEM1	DYSF	EBNA1BP2	EDN2	EEPD1										
	EGFR	EIF2AK3	EIF2B3	EPHA10	EPHA6	EPS15	FABP1	FAM126A	FAM151A										
	FAM176A	FAM183A	FHL3	FOXI3	FOXJ3	FOXO6	GABRR3	GRIK3	GUCA2A										
	GUCA2B	HEYL	HIVEP3	HK2	HPCAL4	hsa-mir-30c-1		hsa-mir-30e	INPP5B										
	KCNQ4	KIAA1841	KIF2C	KRCC1	LRRTM4	MANEAL	MINA	MMACHC	MRPS12	MT-									
ATP8	MTF1	MYCL1	NDUFA4P2	NDUFS5	NFKBIB	NFYC	NT5C1A	OMA1	OR5AC1										
	OR5AC1P	OR5AC2	OR5AC4P	OR5H1	OR5H14	OR5H15	OR5H2	OR5H3P	OR5H4P										
	OR5H5P	OR5H6	OR5H7P	OR5H8P	OR5K3	OR5K4	OSBPL9	OXCT2	PABPC4										
	PAIP2B	PAPOLG	PEX13	PLK3	POLE4	POU5F1P7	PPAP2B	PPIE	PRKAA2										
	PTCH2	PUS10	RASA2	RGPD2	RIMS3	RINL	RNF220	RP11-109P14.1											
	RP11-109P14.10		RP11-109P14.2		RP11-109P14.8		RP11-109P14.9		RP11-124D9.1										
	RP11-15J6.1	RP11-118J21.24		RP11-118J21.5		RP11-191G24.1		RP11-191G24.2											
	RP11-213P13.1		RP11-215L17.1		RP11-223A3.1		RP11-231L11.1		RP11-231L11.3										
	RP11-240D10.2		RP11-240D10.4		RP11-253A20.1		RP11-269F19.2		RP11-269F19.4										
	RP11-275F13.1		RP11-275F13.3		RP11-282K6.3		RP11-316C10.1		RP11-319C21.1										
	RP11-325B23.2		RP11-343D2.11		RP11-346M10.1		RP11-346M10.2		RP11-348A7.1										
	RP11-348A7.2		RP11-377K22.2		RP11-377K22.3		RP11-378I13.1		RP11-399E6.1										
	RP11-399E6.2		RP11-399E6.4		RP11-416L21.1		RP11-438D8.2		RP11-438D8.3										
	RP11-438D8.4		RP11-444F13.3		RP11-444F13.4		RP11-486B10.3		RP11-486B10.4										
	RP11-4H14.1		RP11-526K17.2		RP11-558P9.1		RP11-656D10.3		RP11-656D10.5										
	RP11-656D10.6		RP11-764I5.1		RP11-781D11.1		RP11-789L4.1		RP11-21K4.1	RP1-									
63P18.2	RP4-564O4.1		RP4-580G13.1		RP4-614N24.1		RP4-635E8.1		RP4-657D16.4										
	RP4-657D16.5		RP4-678E16.4		RP4-705F19.1		RP4-705F19.2		RP4-710M16.1										
	RP4-710M16.2		RP4-739H11.1		RP4-739H11.3		RP4-783C10.3		RP5-1066H13.4										
	RP5-1090E8.1		RP5-1103B4.3		RP5-1180C18.1		RP5-860P4.2		RP5-866L20.1										
	RP5-866L20.2		RP5-866L20.3		RP5-882O7.1		RP5-882O7.3		RP5-882O7.4										
	RP6-239D12.1	RPIA	RPL38	RPS8	SARS2	SCMH1	SF3A3	SIRT2											
	SLC25A36	SLFN1	SMYD1	SNORA26	SNORA63	SNORD112	SNORD38	SNORD46	snosnR61										
	snoU109	snoU13	SPSB4	SSBP3	TACR1	TCTEX1D4	TESK2	THNSL2	TMEM53										
	TMEM90B	TRIT1	TTC39A	U2	U5	U6	U7	USP34	UTP11L										
	VAX2	WDR65	Y_RNA	YRDC	ZBTB38	ZNF638	ZNF642	ZNF643	ZNF684										
	ZPLD1																		

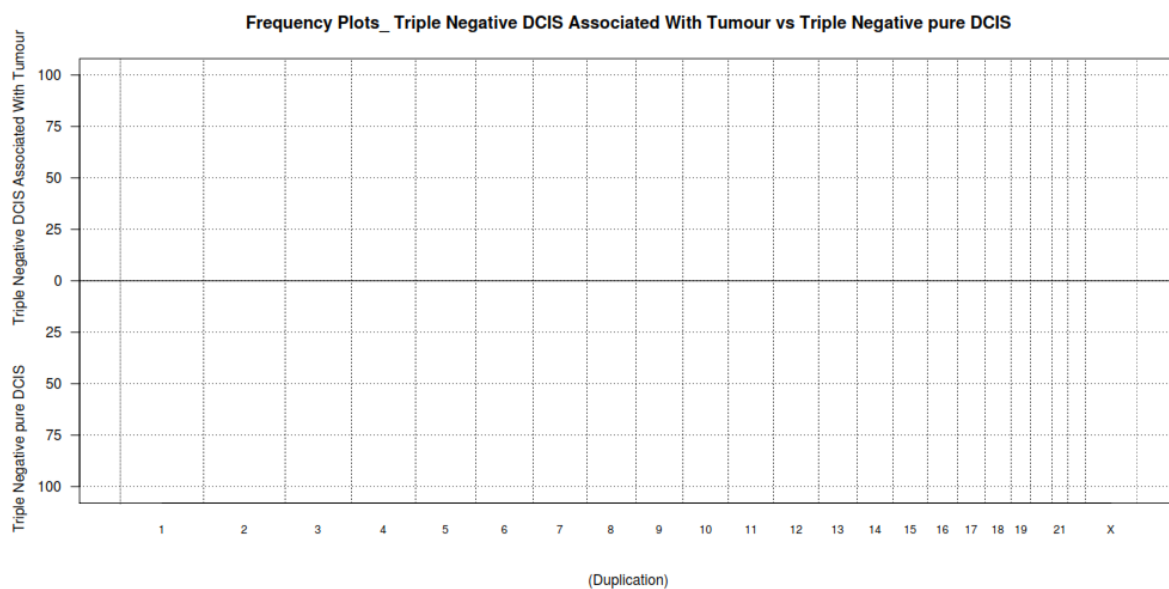
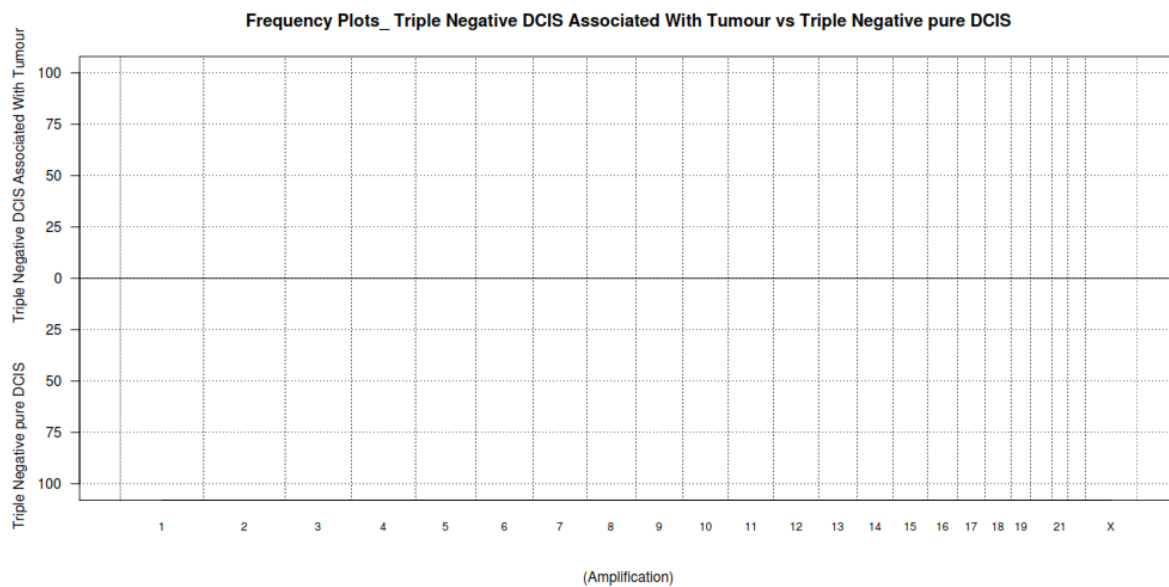
PANTHER analysis: 118 mapped ids are found, 170 mapped ids are not found.

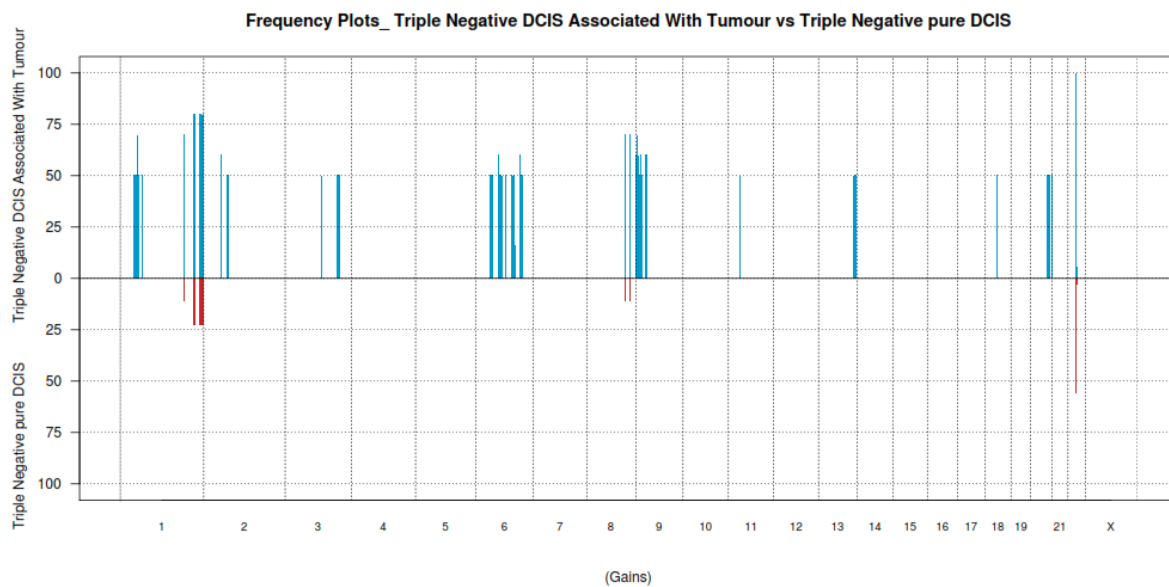
5.8.8 Copy Number Aberrations for Pure Triple Negative DCIS Compared to Triple Negative DCIS Associated with Invasive Breast Disease

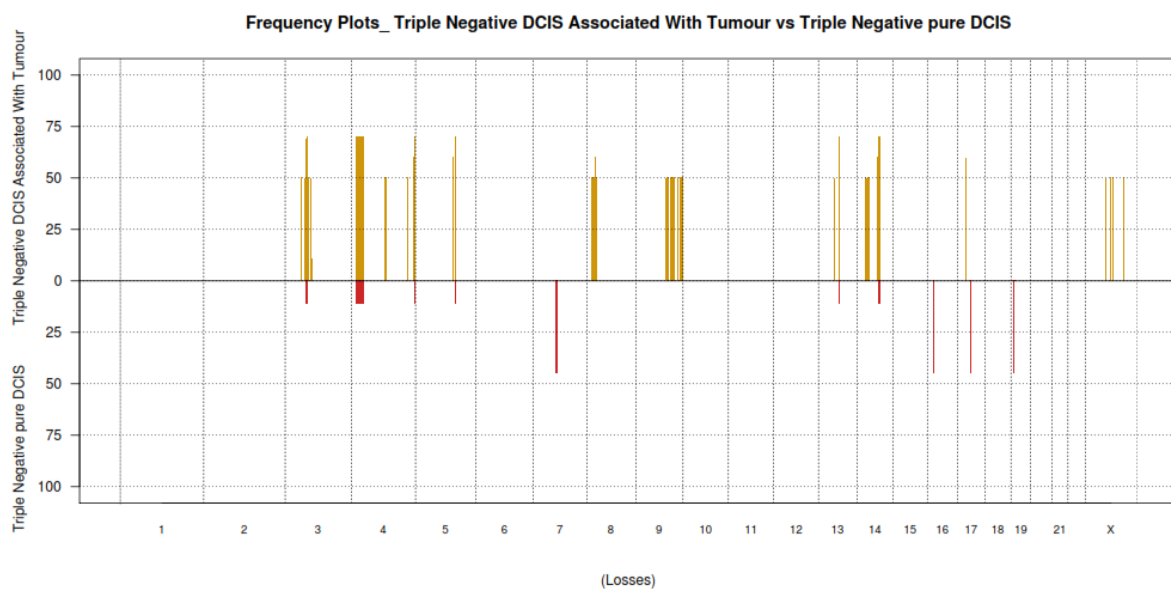
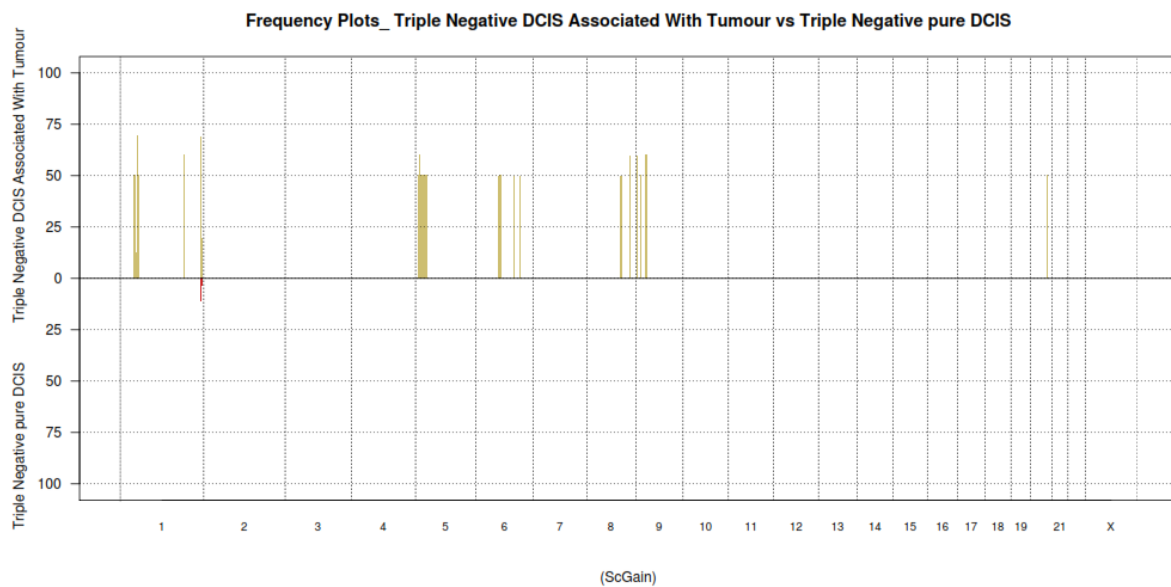
These series examines the differences between pure triple negative DCIS and triple negative DCIS associated with invasive breast disease, i.e. the DCIS component not the invasive breast disease.

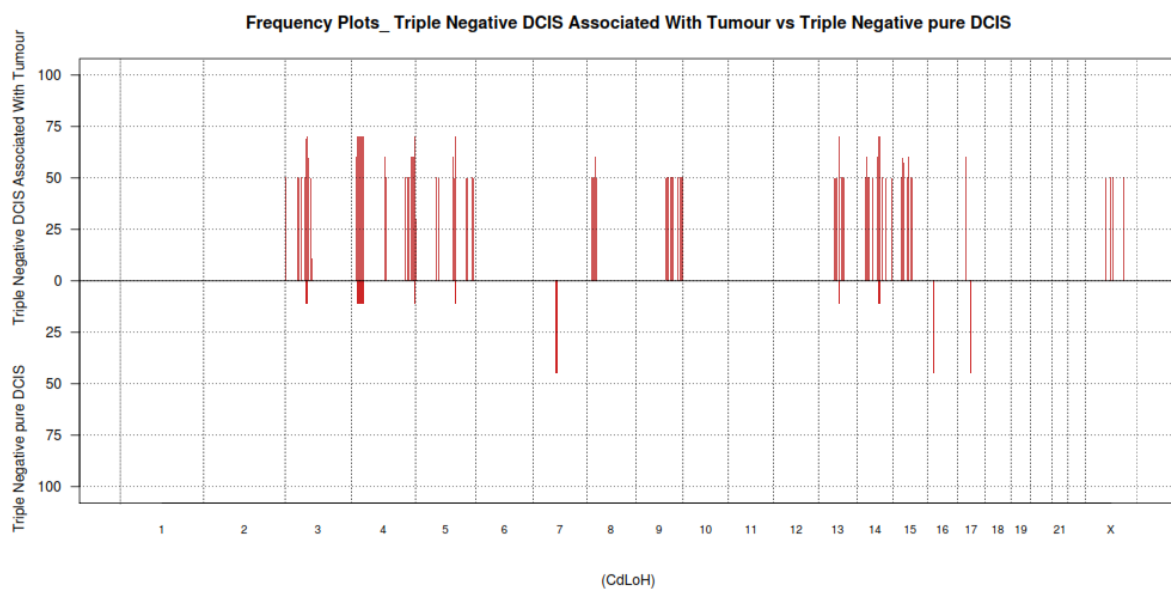
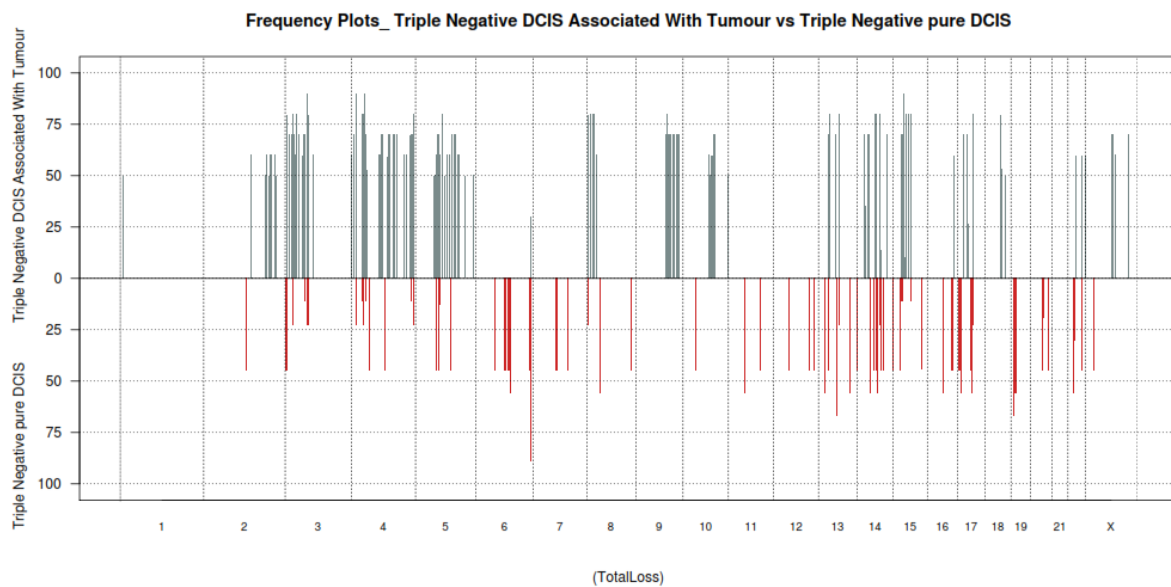
Frequency plots showing copy number aberrations between Pure TN DCIS and TN DCIS associated with invasive breast disease were provided by Breakthrough Breast Cancer/Research Oncology, King's College London Bioinformatics Department (Figure 41).

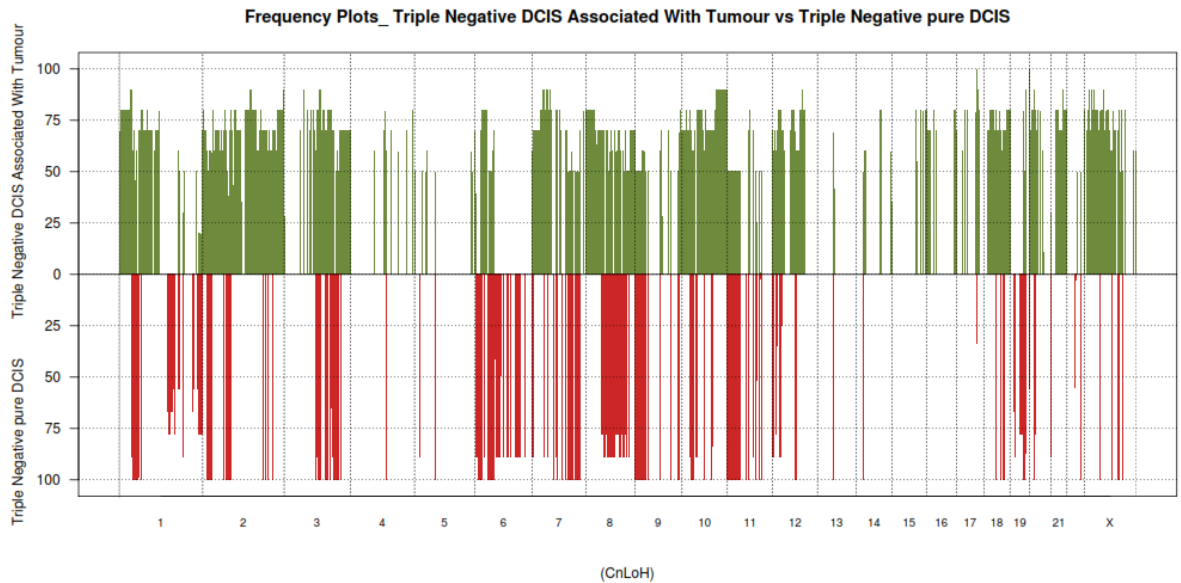
Figure 41: frequency plots showing copy number aberrations between pure triple negative dcis and triple negative DCIS associated with invasive breast disease (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 236-238).











5.8.8.1 Amplification of Pure Triple Negative DCIS and Triple Negative DCIS Associated with Invasive Breast Disease

No amplifications in TN pure DCIS or TN DCIS associated with invasive breast disease were found in this series.

5.8.8.2 Duplication of TN pure DCIS and TN DCIS Associated with Invasive Breast Disease

No Duplications in TN pure DCIS or TN DCIS associated with invasive breast disease found in this series.

5.8.8.3 Genomic Gains of TN pure DCIS and TN DCIS Associated with Invasive breast disease

There is considerable overlap between the genomic gains seen in TN pure DCIS (n=5/9) and TN DCIS associated with invasive breast disease (n=10/10).

1. p-values < 0.05;
2. A gene list is mapped from 13 genomic regions found on chromosomes 1, 8, 22;

- These regions encompass 67 genes altered in TN pure DCIS and TN DCIS associated with invasive breast disease;

Genes:

5S_rRNA	7SK	AC108925.1	ACTN2	AL513363.1	AP000351.1	AP000351.10	AP000351.2	AP000351.6
		AP000351.7	AP000351.8	AP000351.9	CFLP4	CHML	CHRM3	CSMD3
		EDARADD	ENO1P1					
EXO1	FH	GSTT1	GSTTP1	GSTTP2	HEATR1	hsa-mir-1204KMO	KRT18P32	
LGALS8	MAP1LC3C	MTR	NCRNA00210	OPN3	PLD5	PVT1	RP11-149J18.1	
RP11-152L7.1		RP11-152L7.2		RP11-177F15.1	RP11-182B22.2	RP11-307O1.1		
RP11-323D18.1		RP11-323D18.4		RP11-323D18.5	RP11-362H12.1	RP11-371I1.2		
RP11-385F5.2		RP11-385F5.4		RP11-397A15.4	RP11-400N13.1	RP11-439E19.3		
RP11-439E19.5		RP11-439E19.6		RP11-439E19.8	RP11-463J7.2	RP11-488L18.1		
RP11-488L18.8		RP11-544D21.1		RP11-553N16.1	RP11-561I11.2	RP11-561I11.3		
SCCPDH	SNORD112	SPATA17	U6	VN1R17P	VN1R5	WDR64		

PANTHER analysis: 18 mapped ids are found, 49 mapped ids are not found.

There were gains present in TN DCIS associated with invasive breast disease (n=7/7) not observed in pure DCIS (n=0/9).

- p-values < 0.05;
- A gene list is mapped from 13 genomic regions found on chromosomes 1, 2, 6, 9;
- These regions encompass 69 genes altered in TN DCIS associated with invasive breast disease;

Genes:

AC002485.1	AC096583.1	AC139712.2	AL158147.1	AL161450.1	AL161450.10	AL161450.11	AL161450.13	AL161450.16
AL161450.2	AL161450.3	AL161450.4	AL161450.8	AL161450.9	C9orf146	C9orf46	C9orf66	CBWD1
CD274	DMRTA2							
DOCK8	ELAVL4	FAF1	FAM138C	FOXD4	hsa-mir-101-2	hsa-mir-1302-2	JAK2	
LAMA2	NFIB	PDCD1LG2	PTPRD	RCL1	RLN1	RP11-125K10.2	RP11-125K10.4	RP11-125K10.5
RP11-12D24.6	RP11-12D24.7	RP11-143M1.2	RP11-143M1.3	RP11-143M1.4	RP11-143M1.7			
RP11-183G22.1	RP11-183G22.2	RP11-183G22.3	RP11-25O10.1	RP11-280O24.1				
RP11-284P20.2	RP11-284P20.3	RP11-284P20.4	RP11-32F11.2	RP11-39K24.11				
RP11-39K24.14	RP11-39K24.2	RP11-39K24.4	RP11-39K24.5	RP11-39K24.9	RP11-492I2.1			
RP11-560D2.2	RP11-567C20.2	RP11-567C20.3	RP11-572H4.1	RP11-59O6.3	U6			
XXyac-YRM2039.1	XXyac-YRM2039.2	XXyac-YRM2039.3	ZCCHC11					

PANTHER analysis: 16 mapped ids are found, 53 mapped ids are not found.

5.8.8.4 Genomic Sc Gains of TN pure DCIS and TN DCIS Associated with Invasive Breast Disease

There are no significant Sc gains found in pure DCIS in this series. There are Sc gain found in TN DCIS associated with invasive breast disease (n=7/10) that mapped ids in TN pure DCIS (n=0/9).

1. p-values< 0.05;
2. A gene list is mapped from 49 genomic regions found on chromosomes 1, 5, 6, 8, 9, 20;
3. These regions encompass 350 genes altered in TN DCIS associated with invasive breast disease;

Genes:

5S_rRNA	7SK	AC008949.1	AC010455.1	AC010491.1	AC010627.1	AC012640.3	AC016553.1	AC016575.1	AC024589.2
		AC025181.1	AC025472.1	AC025647.3	AC034229.1	AC091878.1	AC099784.2	AC106774.1	AC108925.1
		AC112200.1							
AC114981.1	AC138409.1	AC138409.2	AC138951.3	AC139783.1	ADAMTS12	AGBL4	AL034429.1	AL035467.1	AL133230.2
	AL136164.1	AL136310.1	AL354933.1	AL357519.1	AL390776.1	AL592166.1	AL596225.1	AL645730.1	AL645730.2
AMACR	ANKH	ANKRD33B	ANKRD46	BASP1	BEST4	BTBD19	C1orf123	C1orf163	C1orf185
	C1orf228	C1QTNF3	C5orf17	C5orf22	C5orf23	C6orf155	CC2D1B	CCT5	CDH12
	CDH6	CDH9	CDKN2C	CFLP2	CMBL	COX6CP2	CPT2	CTB-40H15.1	CTB-40H15.2
	CTB-40H15.2		CTB-40H15.3		CTB-55B8.1	CTC-305H11.1		CTC-461F20.1	
	CTD-2001E22.1		CTD-2001E22.2	CTD-2004A9.1		CTD-2008E3.1		CTD-2010I22.1	
	CTD-2010I22.2	CTD-2011G10.1		CTD-2011G17.1		CTD-2017D15.1		CTD-2023M8.1	
	CTD-2057J6.1		CTD-2057J6.2		CTD-2061E9.1		CTD-2066L21.1	CTD-2066L21.2	
CTD-2066L21.3		CTD-2134P3.1		CTD-2138O14.1		CTD-2139B15.1	CTD-2139B15.2	CTD-2139B15.4	
	CTD-2139B15.5		CTD-2143L24.1		CTD-2151L9.1		CTD-2151L9.2	CTD-2154B17.1	
	CTD-2165H16.1		CTD-2165H16.2		CTD-2165H16.3		CTD-2194F12.1	CTD-2197O4.1	
	CTD-2199O4.1		CTD-2201E9.1		CTD-2201E9.2		CTD-2201E9.3	CTD-2201E9.4	
	CTD-2203K17.1		CTD-2206G10.1		CTD-2206G10.2		CTD-2215L10.1	CTD-2218G20.1	
	CTD-2218G20.2		CTD-2233C11.2		CTD-2233C11.3		CTD-2256P15.1	CTD-2269E23.1	
	CTD-2269E23.2		CTD-2290C23.1		CTD-2290C23.2		CTD-2290C23.3	CTD-2290P7.1	
	CTD-2313D3.1		CTD-2351A8.1		CTD-2533K21.1		CTD-2533K21.2	CTD-2533K21.3	
	CTD-2533K21.4		CTD-3007L5.1		CTD-3065A13.1		CTNND2	DAP	DMRTA2
	DMRTB1	ECHDC2	EIF2B3	ELAVL4	ERI3	FAF1	FAM105A	FAM105B	FAM134B
	FAM173B	FBXL7	FBXO43	FTHL10	GPX7	GRHL2	GUCA2A	GUCA2B	HIVEP3
mir-30a	hsa-mir-30c-2		hsa-mir-579	KB-1615E4.1	KIF2C	LAMA2	LRP8	MAGOH	Mar11
	MTMR12	MYCL1	MYO10	NCALD	NPR3	OGFRL1	ORC1L	PDZD2	PLK3
	PMCHL1	PRDM9	PRPF38A	PTCH2	PTPN1	REV3L	RGS22	RIMS1	RNASEN
	RNF19A	RNF220	ROPN1L	RP11-1023L17.1	RP11-1023L17.2		RP11-1084J3.1		RP11-1084J3.2
RP11-1084J3.3		RP11-113I22.1		RP11-117D22.1		RP11-117D22.2		RP11-124N3.1	
	RP11-124N3.2		RP11-124N3.3		RP11-1250I15.1		RP11-1250I15.2		RP11-1325J9.1
	RP11-154D6.1		RP11-155O18.3		RP11-159C21.3		RP11-159C21.4		RP11-183G22.1

PANTHER analysis: 6 mapped ids are found, 23 mapped ids are not found.

There are some similarities in genomic losses in TN DCIS associated with invasive breast disease (n=7/10) compared to TN pure DCIS (n=1/9).

1. p-values < 0.05;
2. A gene list is mapped from 19 genomic regions found on chromosomes 3, 4, 5, 13, 14;
3. These regions encompass 49 genes altered in TN DCIS associated with invasive breast disease;

Genes:

AC006427.1	AC006427.2	AC006427.3	AC093848.1	AC106868.1	AC107394.1	AC109351.1	AC110766.1	AL590102.1	AL596329.1
CCKAR	CTC-507E12.1	FHIT	GPHB5	GPHN	KCNH5	KCNIP4	LDB2	MAGI1	PACRGL
PCDH7	PPP2R5E	RHOJ	RP11-120A1.1		RP11-141E13.1		RP11-174E22.1		RP11-
174E22.2	RP11-293A21.1	RP11-293A21.2		RP11-308K2.1		RP11-315A17.1		RP11-390C19.1	
	RP11-415C15.1	RP11-417M17.1		RP11-429G17.1		RP11-495L18.2		RP11-543A18.1	
	RP11-617I14.1	RP11-619J20.1		RP11-93M12.1		SCARNA20	SLIT2	SNORD74	
snoU13	STIM2	SYNE2	TAPT1	TBC1D19	U6				

PANTHER analysis: 16 mapped ids are found, 33 mapped ids are not found.

5.8.8.6 TN DCIS Associated with Invasive breast disease

There were total losses present in TN pure DCIS (n=8/9) which were not observed in TN DCIS associated with invasive breast disease (n=0/10).

1. p-value<0.05;
2. A gene list is mapped from 25 genomic regions found on chromosomes 6, 8, 11, 13, 14, 16, 17, 19, 22;
3. These regions encompass 72 genes altered in TN pure DCIS;

Genes:

AARSD1	ABCC12	ABCC4	AC004019.10	AC004019.17	AC004659.1	AC005513.1	AC008734.1
AC008734.2	AC087650.1	AC087650.8	AC090427.1	AC100793.2	ADAM5P	AL136359.1	AL137001.1
AOC2	AOC3						
ARL4D	BRCA1	CACNA1A	CCDC105	CECR2	CTD-3199J23.2	DNAH9	EMR2
EMR3	FBXO33						
G6PC	GPC5	hsa-mir-2117	IFI35	MBD3L1	MUC16	OLFM4	OR4C3
OR4C45	OR4C5	OR4S1					
OR4X1	OR4X2	OR7A10	OR7A11P	OR7A17	OR7A2P	OR7A5	OR7C1
OR7C2	PARK2	PSME3					
RND2	RP11-24H2.1	RP11-24H2.2		RP11-301J16.2	RP11-301J16.3		RP11-301J16.5

RP11-301J16.7 RP11-442J17.4 RPL27 RTN1 RUNDC1 SLC1A6 SNORA40 U2 U6
VAT1 Y_RNA ZNF333 ZNF558

PANTHER analysis: 40 mapped ids are found, 29 mapped ids are not found.

There were total losses present in TN DCIS associated with invasive breast disease (n=8/10) which were not observed in TN pure DCIS (n=0/9).

1. p-values < 0.05;
2. A gene list is mapped from 32 genomic regions found on chromosomes 3, 4, 5, 8, 9, 13, 14, 15, 17, 18;
3. These regions encompass 56 genes altered in TN DCIS associated with invasive breast disease;

Genes:

7SK	AC021860.1	AC022559.1	AC023385.1	AC034111.1	AC084882.1	AC091133.1	AC096750.1	AC103702.1	AC114477.1
	AC139426.1	AC139426.3	AC139426.4	ACTR10	AL079307.3	AL079307.4	AL079307.5	AL121579.1	AL132989.2
AL138995.1	AL139021.2	AL163952.2	ARAP2	ARID4A	ASPN	B4GALNT2	C14orf166	C14orf181	CENPP
CSMD1	CTD-2015H6.1			CTD-2015H6.2			CTD-2015H6.3	CTD-2244F11.2	
	CTD-2325P2.2		DTHD1	ECM2	FAM151B	FAM7A1	FMN1	FRMD6	GIP
GNG2	GREM1	HMGB1L14	HOXB7	HOXB9	IGF2BP1	KIAA0586	KLB	KLF3	KLHL5
MITF	MSRA	MTUS1	OMD	PDCD6IP	PDS5B	PELI2	PRSS55	PSMA3	
RAD51L1	RFC1	RP11-1000B6.2		RP11-212F11.1	RP11-213G21.1		RP11-213G21.2		
RP11-22A3.2		RP11-360F5.1		RP11-360F5.2	RP11-360F5.3		RP11-380B4.2		
RP11-431M7.2		RP11-431M7.3		RP11-501C14.5		RP11-501C14.6		RP11-501C14.7	
RP11-617D20.1		RP11-708H21.1		RP11-83C7.1		RP11-83C7.2		RP13-395E19.1	
RP1L1	RPL9	SCG5	SNF8	SNORA18	snoU13	SPRED1	TIMM9	TMEM156	TOMM20L
U3	U6	U8	UBE2Z	WDR19	Y_RNA	ZFP36L1	ZFYVE16		

5.8.8.7 CdLOH of TN pure DCIS and TN DCIS Associated with Invasive breast disease

There is CdLOH present in TN pure DCIS (n=4/9) which were not observed in TN DCIS associated with invasive breast disease (n=0/10)

1. p-values < 0.05;
2. A gene list is mapped from 4 genomic regions found on chromosomes 7, 16, 17;

3. These regions encompass 12 genes altered in TN pure DCIS;

Genes:

AC006013.1 AC131384.1 CALN1 KRT14 KRT16 KRT17 KRT42P RP11-358M3.1 RP4-635O5.1 RP5-945F2.1 RP5-945F2.2 RP5-945F2.3

PANTHER analysis: 4 mapped ids are found, 8 mapped ids are not found.

There is CdLOH present in TN DCIS associated with invasive disease (n=7/10) not observed in TN pure DCIS (n=0/9).

1. p-values < 0.05;
2. A gene list is mapped from 37 genomic regions found on chromosomes 3, 4, 5, 8, 13, 14, 15, 17;
3. These regions encompass 96 genes altered in TN DCIS associated with invasive breast disease;

Genes:

7SK AC006160.1 AC006160.2 AC006160.4 AC006160.5 AC006445.6 AC006445.7 AC007126.1 AC015688.1 AC020698.1
AC093848.1 AC098830.1 AC106868.1 AC107394.1 AC110766.1 AC111152.1 ADH1A ADH1B ADH1C
AL110292.1 BOD1L C14orf53 C4orf47 C5orf13 CCDC110 CDCA2 CLRN2 CTC-459M5.1 CTC-459M5.2
CTD-2022H16.1 CTD-2192A1.1 EPB41L4A FAM184B FAM189A1 FAM19A4 GBA3 GPM6A
GPR125 KCNH5 KCNIP4 KSR1 LAP3 LDB2 LGALS9 MED28 MTNR1A
NCRNA00219 PCDH7 PDLIM3 PPARGC1A PTPRG QDPR RP11-103J17.1 RP11-10G12.1
RP11-10G12.2 RP11-120A1.1 RP11-141E13.1 RP11-17E2.2 RP11-196P2.1
RP11-215A19.1 RP11-24D11.1 RP11-279O9.4 RP11-301L8.2 RP11-315A17.1
RP11-333E5.1 RP11-333E5.2 RP11-341G5.1 RP11-380P13.1 RP11-380P13.2
RP11-495L18.2 RP11-540E16.2 RP11-556G22.1 RP11-556G22.2 RP11-556G22.3
RP11-576E20.1 RP11-598O12.1 RP11-598O12.2 RP11-617I14.1 RP11-619J20.1
RP11-626E13.1 RP11-665I14.1 RP11-696N14.1 RP11-722M1.1 RP11-806K15.1
RP11-93M12.1 RP13-497K6.1 RPL10AP6 SNORA13 SNORA51 SNORA75 snoU13 SORBS2
TJP1 U6 WDR17 Y_RNA

PANTHER analysis: 28 mapped ids are found, 68 mapped ids are not found.

5.8.8.8 CnLOH of TN pure DCIS and TN DCIS Associated with Invasive Breast Disease

There is CnLOH present in TN pure DCIS (n=9/9) which were not observed in TN DCIS associated with invasive breast disease (n=0/10).

- Genes:

248

397G17.1	RP1-13D10.2	RP1-13D10.3	RP1-13D10.4	RP1-13D10.5	RP11-401E14.1
	RP11-401E14.2	RP11-411K7.1	RP11-411K7.2	RP11-411K7.4	RP11-411K7.5
	RP11-418B12.1	RP11-420A21.1	RP1-142O9.2	RP11-430C17.1	RP11-446F17.3
	RP11-451G4.2	RP11-451G4.3	RP11-460H18.1	RP11-460H18.2	RP11-460N20.3
	RP11-460N20.4	RP11-460N20.5	RP11-462L8.1	RP11-462L8.2	RP11-466L17.1
RP11-470H9.1	RP11-472N13.2	RP11-479G22.3	RP11-479G22.5	RP11-479G22.6	
	RP11-479G22.7	RP11-501I19.3	RP11-501I19.4	RP11-501O2.1	RP11-501O2.2
	RP11-501O2.3	RP11-501O2.4	RP11-501O2.5	RP11-505J9.1	RP1-151F17.1
	RP11-529G21.4	RP1-152L7.1	RP1-152L7.3	RP1-152L7.5	RP11-548O1.1
					RP11-548O1.3
RP11-550C4.6	RP11-590C23.1	RP11-598F17.1	RP11-630I5.1	RP11-635I10.1	RP11-637O19.2
	RP11-639B1.1	RP11-63E16.1	RP11-649A16.1	RP11-649A16.5	RP11-651P23.2
	RP11-651P23.3	RP11-651P23.4	RP11-651P23.5	RP11-651P23.6	RP11-666A20.1
	RP11-666A20.3	RP11-666A20.4	RP11-689D3.4	RP11-71N10.1	RP11-722C17.1
	RP11-723O4.2	RP11-735A5.1	RP11-746P2.5	RP11-767N6.2	RP11-767N6.7
	RP11-771D21.2	RP11-788A4.1	RP11-788A4.2	RP11-793K1.1	RP11-795J1.1
	RP11-795J1.2	RP11-797H7.1	RP11-797H7.2	RP11-797H7.3	RP11-797H7.5
	RP11-79N23.1	RP11-7O11.3	RP1-180E22.3	RP11-812I20.2	RP11-90C4.1
	RP11-90C4.2	RP11-90C4.3	RP11-90C4.4	RP1-190J20.2	RP11-91A18.1
	RP11-91A18.4	RP11-91A18.5	RP11-9N20.1	RP1-202I21.3	RP1-202I21.5
	RP1-207H1.2	RP1-207H1.3	RP1-217P22.2	RP1-228H13.1	RP1-228H13.2
257A7.4	RP1-273P12.1	RP1-273P12.3	RP1-27K12.2	RP1-27K12.4	RP1-290I10.4
290I10.5	RP1-290I10.7	RP1-295F6.2	RP1-304B14.3	RP1-319M7.2	RP13-469O16.1
	RP13-476E20.1	RP1-62D2.3	RP1-71H19.2	RP3-322I12.2	RP3-322L4.2
					RP3-335N17.2
	RP3-347E1.2	RP3-365O12.2	RP3-380E11.2	RP3-437C15.1	RP3-445N2.1
					RP3-
448I9.1	RP3-448I9.2	RP3-451B15.3	RP3-462C17.1	RP3-486B10.1	RP3-486D24.1
					RP3-503A6.2
	RP3-510L9.1	RP4-533D7.3	RP4-533D7.4	RP4-533D7.5	RP4-633H17.2
					RP4-
657D16.3	RP4-724P12.1	RP4-753D5.3	RP4-753D5.4	RP4-761I2.2	RP4-761I2.4
					RP5-1043E3.1
	RP5-864K19.4	RP5-994D16.3	RP5-994D16.7	RP5-994D16.9	RPL15P3
					RPN1
	RPS15AP10	RPS15AP16	RPS17P5	RPS6KA4	RRAGC
					SEC22A
					SFRS16
					SHKBP1
					SHMT1P1
	SIRT5	SLC25A20P1	SLC25A27	SLC35B3	SLC6A9
					SMAP2
					SNORA15
					SNORA22
	SNORA24	SNORA5	SNORA8	SNORD45	SNORD66
					snoU13
					SOX4
					SPTBN4
					ST3GAL3
					STX1A
	SUCLA2P2	SYCP2L	TBC1D7	TBL2	TDRD6
					TFAP2B
					TFAP2D
					TFEC
					THAP5
					TMEM125
	TMEM14A	TMEM14B	TMEM14C	TMEM69	TMSB10
					TNFRSF21
					TPMT
					TRAM2
					TRAPPC6A
	TSPAN1	TXNDC12	U1	U3	U6
					U7
					USP34
					VPS37D
					WBSCR22
					WDR92
	WWTR1	Y_RNA	YBX1	ZEB1	ZFAND3
					ZIC1
					ZIC4
					ZMYND12
					ZNF117
	ZNF138	ZNF273	ZNF296	ZNF438	ZNF643
					ZNF691

PANTHER analysis: 126 mapped ids are found, 331 mapped ids are not found.

There is some overlap in CnLOH in TN DCIS associated with invasive breast cancer (n=10/10) and CnLOH in TN pure DCIS (n=5/9).

This represents CnLOH for TN DCIS associated with invasive breast disease.

1. p-values < 0.05;

2. A gene list is mapped from 49 genomic regions found on chromosomes 1, 2, 3, 7,10,12,17, 19, 20, 21, X;
3. These regions encompass 309 genes altered in TN DCIS associated with invasive breast disease;

Genes:

5S_rRNA	AC004455.1	AC004691.5	AC005189.6	AC005537.2	AC005879.1	AC006466.5	AC006960.7	AC006978.1	AC010789.1										
	AC015923.1	AC016706.1	AC063927.2	AC063927.3	AC063927.4	AC063927.5	AC063927.6	AC063927.7	AC063927.8										
	AC063927.9	AC063944.1	AC068765.1	AC073585.1	AC087069.2	AC113554.1	AC114788.1	AC114788.2	AC138128.1										
ACADSB	ACE	ACE2	ADCYAP1R1	ADRB1	AF015262.1	AF015262.2	AF015720.3	AF015726.1											
	AF020802.1	AF020802.2	AF020802.3	AF020802.4	AL133215.1	AL133387.2	AL135794.1	AL157788.1	AL158014.1										
	AL162407.1	AL353664.1	AL355598.1	AL355861.1	AL359747.1	AL359836.1	AL591845.1	AL603764.1	AL731543.1										
ANLN	AOAH	AP000688.8	AQP1	ARMS2	ATE1	BBX	BLVRA	BMX											
BRIP1	BTBD16	BTRC	BUB3	C10orf118	C10orf119	C10orf120	C10orf2	C10orf46											
	C10orf76	C10orf84	C10orf88	C10orf90	C1orf113	C7orf44	CA5B	CA5BP	CASC2										
CCDC54	CD3EAP	CSF3R	CUZD1	CYB561	DCAF7	DCLRE1A	DLGAP3	DMBT1											
DNMBP	DOCK1	DPCD	EIF3A	ELOVL3	EMX2	EMX2OS	ERCC1	FAM176B											
FAM178A	FAM188B	FAM24A	FAM24B	FAM45A	FBXW4	FGF8	FGFR2	FOSB											
FTLP18	GBF1	GHRHR	GLI3	GPAM	GRK5	GS1-594A7.3		GS1-594A7.5											
HDAC4	HECW1	HMX2	HMX3	HPS6	hsa-mir-608	hsa-mir-802	HTRA1	IKZF5											
INMT	INPP5F	INTS2	JAKMIP3	KAZALD1	KCNH6	KCNIP2	KCNK18	KIAA1598											
LBX1	LDB1	LSM10	LZTS2	MGEA5	MRPL43	MRPS15	MRPS24	MT-ATP6	MT-										
CO1	MT-CO3	MT-ND1	MT-ND2	MT-ND3	MT-ND4	MT-ND4L	MT-ND5	MT-ND6	NANOS1										
	NCRNA00093	NHLRC2	NKX2-3	NOLC1	NPM3	NSMCE4A	OSCP1	PDZD7											
	PDZD8	PITX3	PLEKHA1	POLL	PPM1N	PPP1R13L	PPP1R2P2	PPRC1	PRDX3										
	PRLHR	PSTK	RAB11FIP2	RGS9	RP11-107C16.2		RP11-107I14.1	RP11-107I14.2											
	RP11-107I14.4		RP11-107I14.5		RP11-108L7.4		RP11-108L7.7	RP11-129J12.1											
	RP11-153N17.1		RP11-162A23.5		RP11-179H18.8		RP11-181E22.1	RP11-190J1.10											
	RP11-190J1.3		RP11-190J1.7		RP11-198M6.2		RP11-198M6.5	RP11-214N15.5											
	RP11-215A21.2		RP11-223P11.2		RP11-223P11.3		RP11-244H3.1	RP11-244H3.2											
	RP11-245J24.1		RP11-248I9.2		RP11-25C19.1		RP11-25C19.3	RP11-280F2.1											
	RP11-280F2.2		RP11-282I1.1		RP11-298H24.1		RP11-302K17.3	RP11-302K17.4											
	RP11-302K17.7		RP11-309P22.1		RP11-319I23.2		RP11-319I23.3	RP11-31L23.3											
	RP11-324L3.1		RP11-328K15.1		RP11-338O1.2		RP11-354M20.3	RP11-36N22.1											
	RP11-36N22.3		RP11-381K7.1		RP11-381K7.2		RP11-411P18.2	RP11-446H18.1											
	RP11-446H18.3		RP11-446H18.5		RP11-45P22.1		RP11-45P22.2	RP11-462G8.2											
	RP11-462G8.3		RP11-498J9.1		RP11-498J9.2		RP11-498J9.4	RP11-500G22.2											
	RP11-501J20.2		RP11-501J20.3		RP11-501J20.5		RP11-561O4.1	RP11-564D11.3											
	RP11-567J24.4		RP11-56I23.1		RP11-57H14.3		RP11-5G18.2	RP11-781P14.3											
	RP11-78C6.1		RP11-79M19.2		RP11-93B21.1		RP13-314C10.5	RP5-1158B12.1											
	RP5-877J2.1	RPL23AP3	RPL34P3	RPS20P1	RTN2	RUNX1	SAT1	SCN4A	SEC23IP										
	SEMA4G	SETD4	SFPQ	SFXN3	SFXN4	SLC18A2	SNORA19	SNORA63	SNORD112										
	SNORD60	snoU13	SORCS3	STK40	TACC2	TANC2	TCERG1L	THRAP3	TLX1										
	TLX1NB	TMEM27	TWIST2	U2	U3	U4	U6	URGCP	USP29										
	VASP	VAX1	VT11A	Y_RNA	ZIM3	ZMYM1	ZMYM6												

PANTHER analysis: 134 mapped ids are found, 157 mapped ids are not found.

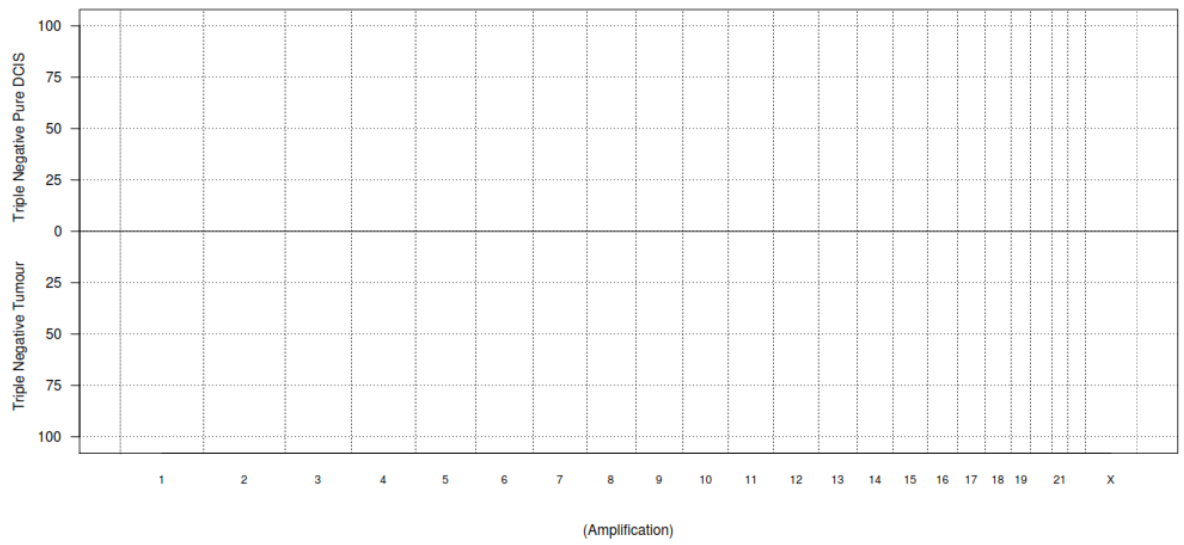
5.8.9 Copy Number Aberrations for Triple Negative Pure DCIS Compared to Triple Negative Invasive Breast Disease

These series examines the difference between pure triple negative DCIS only and TN invasive breast disease.

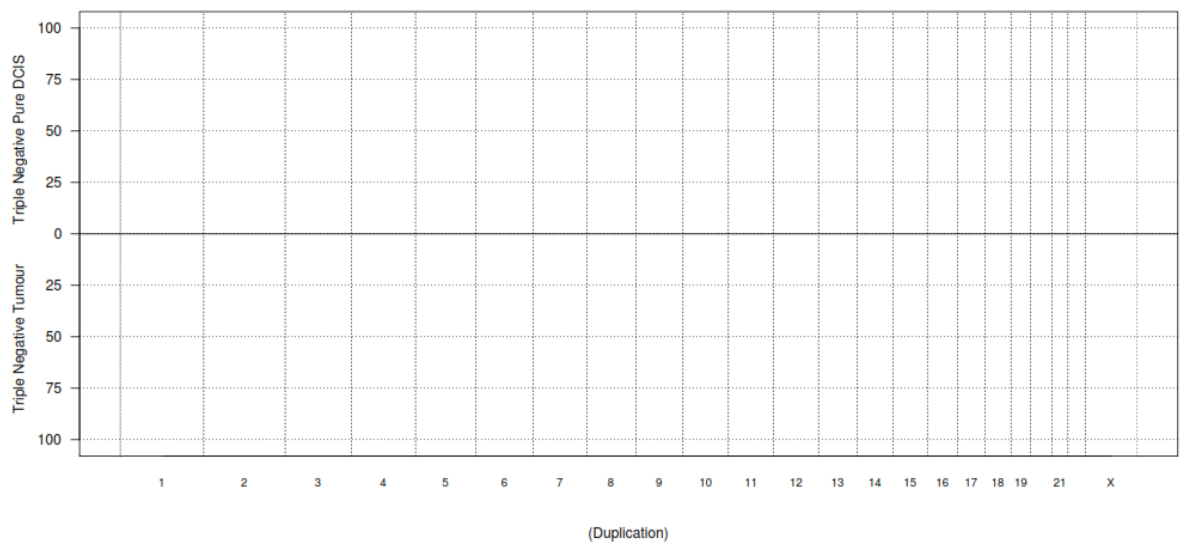
Frequency plots showing copy number aberrations between Pure TN DCIS and TN invasive breast disease were provided by Breakthrough Breast Cancer/Research Oncology, King's College London Bioinformatics Department (Figure 42).

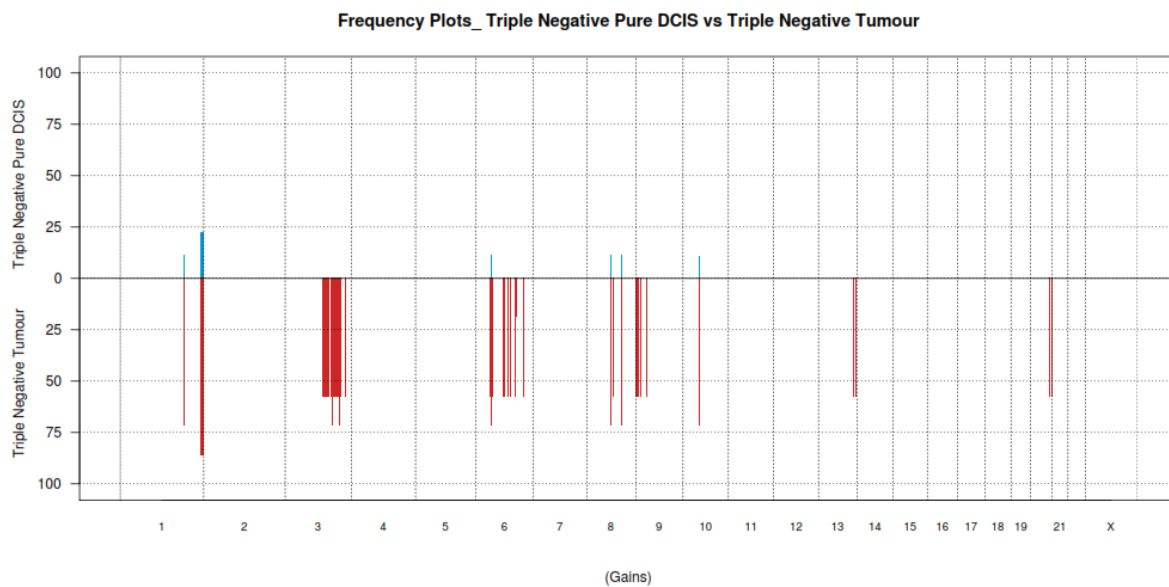
Figure 42: Frequency plots showing copy number aberrations between pure triple negative DCIS and triple negative invasive breast disease (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 249-251).

Frequency Plots_ Triple Negative Pure DCIS vs Triple Negative Tumour

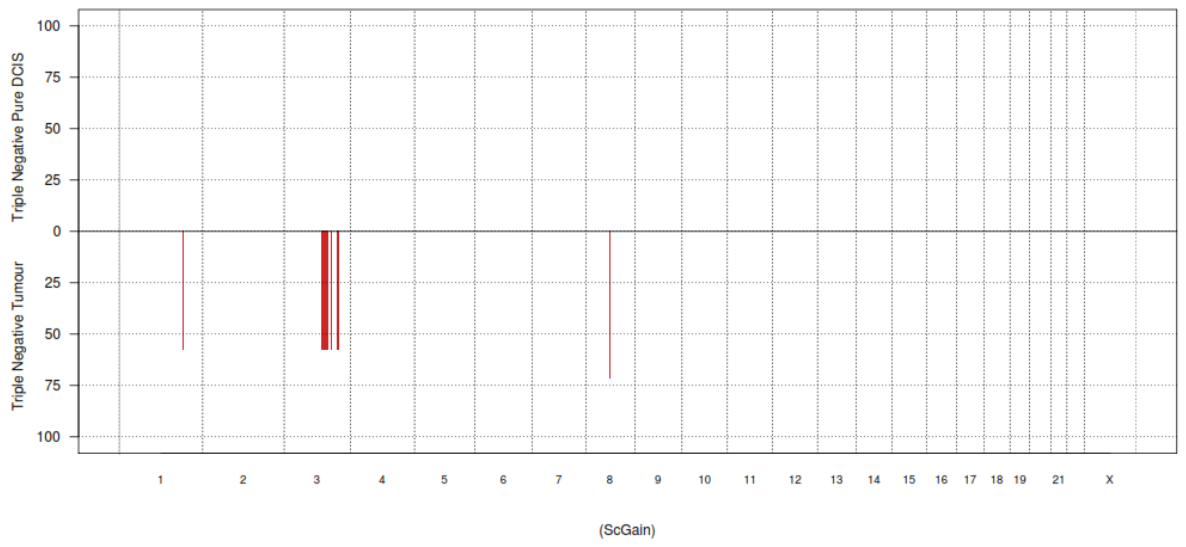


Frequency Plots_ Triple Negative Pure DCIS vs Triple Negative Tumour

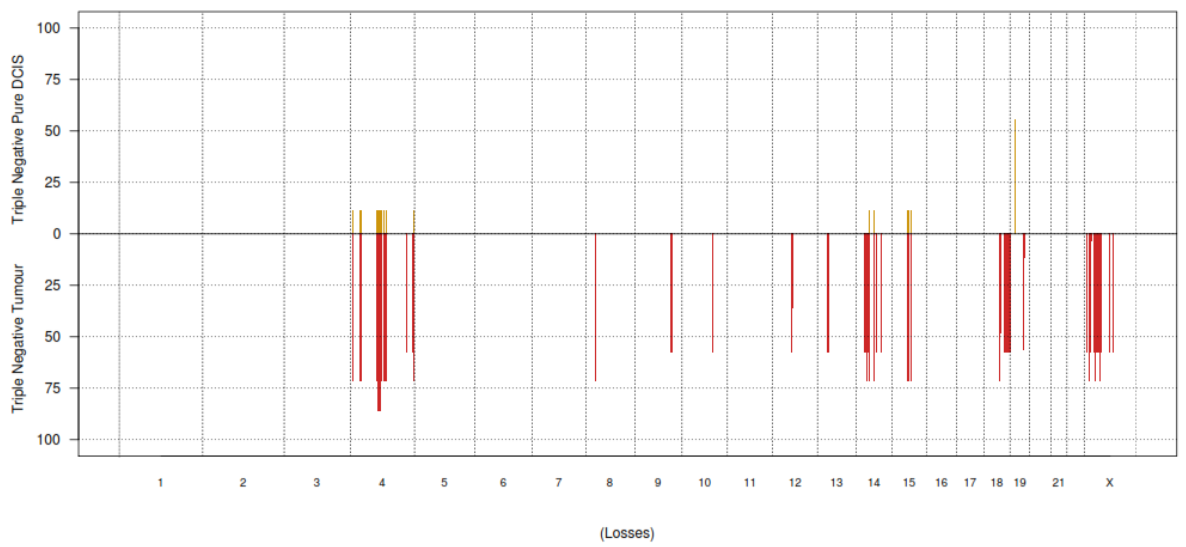




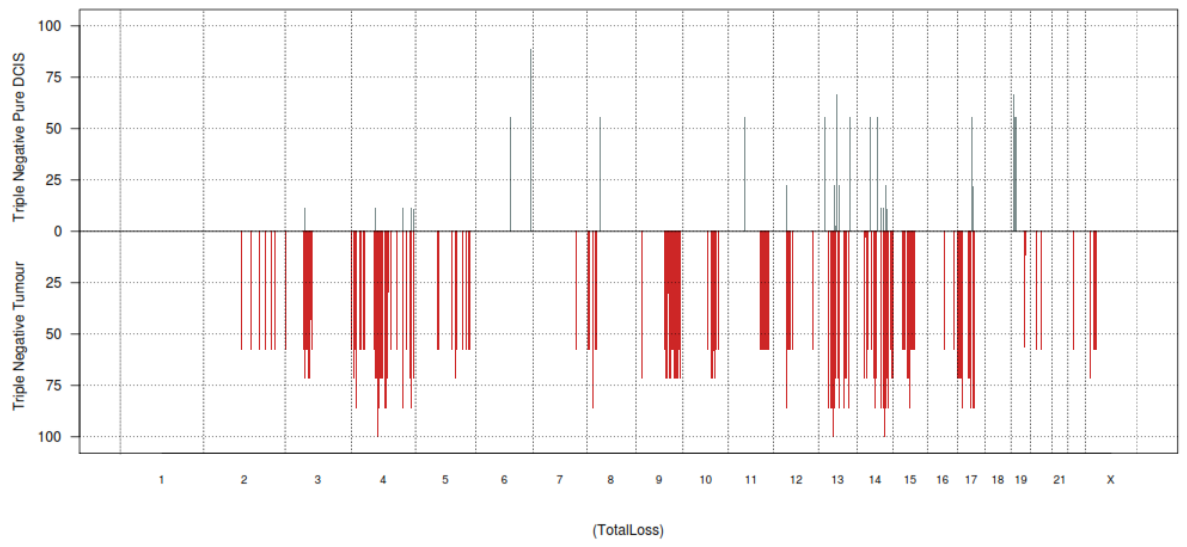
Frequency Plots_ Triple Negative Pure DCIS vs Triple Negative Tumour



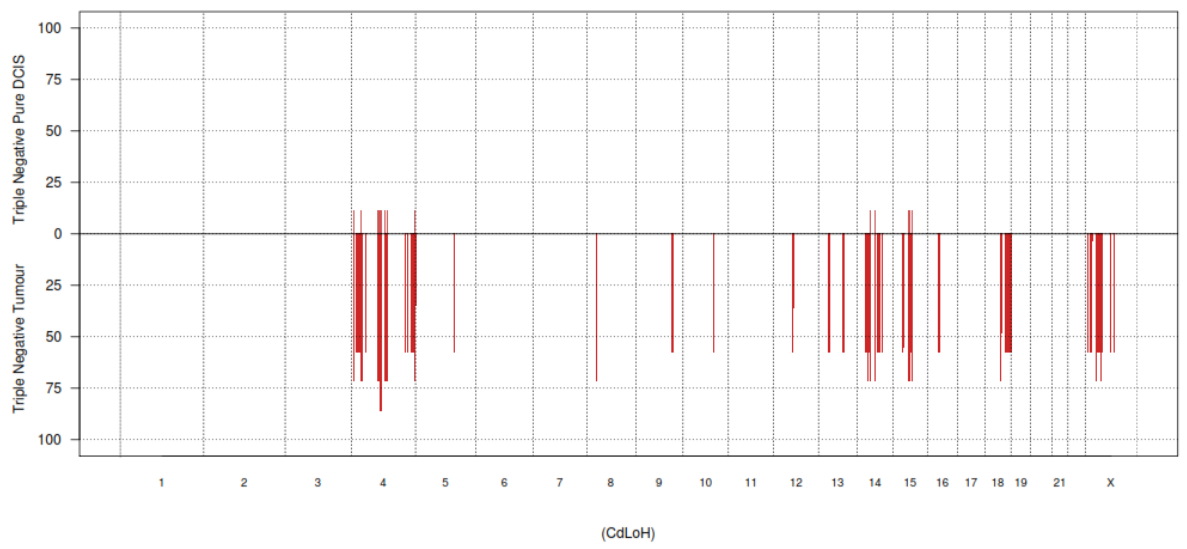
Frequency Plots_ Triple Negative Pure DCIS vs Triple Negative Tumour

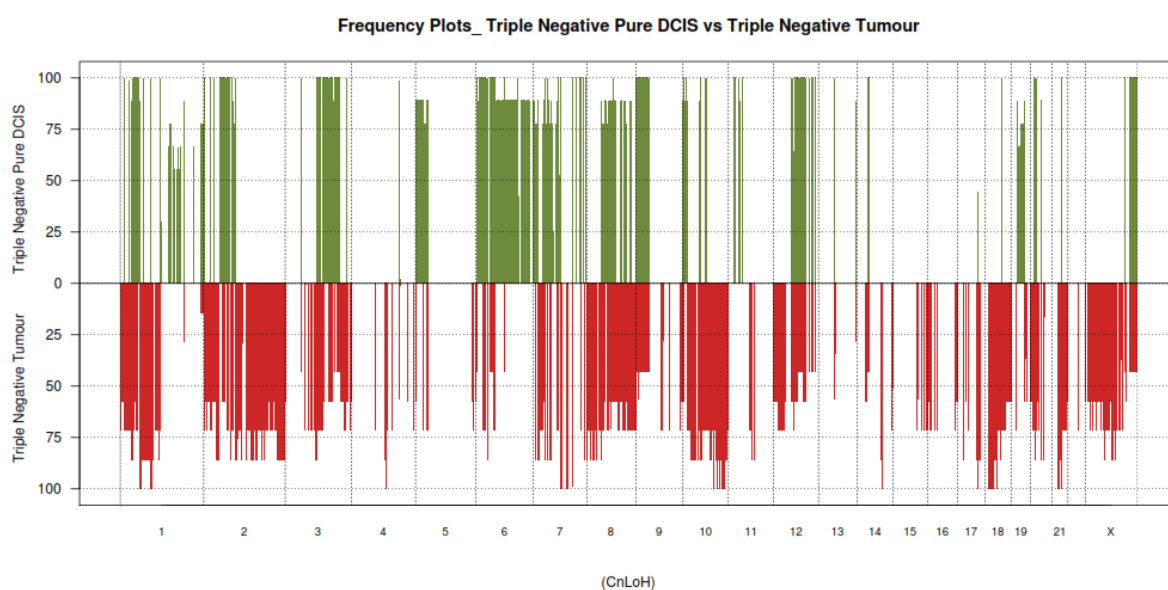


Frequency Plots_ Triple Negative Pure DCIS vs Triple Negative Tumour



Frequency Plots_ Triple Negative Pure DCIS vs Triple Negative Tumour





5.8.9.1 Amplification of Pure Triple Negative DCIS and Triple Negative Invasive Breast Disease

No amplifications are found in TN pure DCIS or TN invasive breast disease in this series.

5.8.9.2 Duplication in Pure Triple Negative DCIS and Triple Negative Invasive Breast Disease

There are duplications are found in TN pure DCIS (n=1/9) and TN invasive breast disease (n=3/7).

1. p-values < 0.05;
2. A gene list is mapped from 13 genomic regions found on chromosomes 2, 6, 8, 12;
3. These regions encompass 181 genes altered in TN pure DCIS and TN invasive breast disease;

Genes:

5S_rRNA	7SK	ABCC10	AC012494.1	AC079600.1	AL080315.3	AL096711.1	AL121959.1	AL136304.1	AL136304.2
AL136304.3	AL451073.1	AL583834.1	ARG1	BMP5	BYSL	C6orf108	C6orf132	C6orf15_2	C6orf154
C6orf174	C6orf192	C6orf226	C6orf58	C8orf85	CCND3	CENPW	CLIC5	CNPY3	COL21A1
CTAGE9	CTGF	CUL7	CUL9	DLK2	ECHDC1	EEF1A1P36	EEF1DP5	ENPP1	ENPP3

ENPP4	ENPP5 EYA4 FRS3	GFRAL	GNMT	GUCA1A	GUCA1B	HCRT2	HEY2	HINT3
HMGCL1L13	HMGCLL1	hsa-mir-588	KIAA0240	KLC4	KLHDC3	LAMA2 LRFN2	MEA1	MED20
MED23	MOXD1	MRPL2	MRPS10 NCOA7 OR2A4	PEX6	PGC	POLR1C	PPP2R5D	PRB2
PRICKLE1	PRICKLE4 PTCRA	PTK7	PTPRK	RNF146	RP11-103C16.2		RP11-121P10.1	RP11-
123H21.1	RP11-151M7.1	RP11-162L10.1		RP11-203B4.1	RP11-213N20.1		RP11-228O6.2	
RP11-251I15.1	RP11-295F4.4	RP11-298J23.5	RP11-298J23.6		RP11-314E23.1		RP1-131F15.2	
RP11-325O24.1	RP11-325O24.2	RP11-357P24.2	RP1-139D8.6		RP11-480N24.3		RP11-480N24.4	
RP11-480N24.6	RP11-527F13.1	RP11-527F13.2		RP11-533O20.2		RP11-570K4.1		
RP11-624M8.1	RP11-691I8.2	RP11-691I8.3	RP11-753G20.1	RP11-775I22.2	RP1-177A13.1		RP1-179E13.1	
RP1-283K11.3	RP1-293L8.2	RP1-293L8.5	RP13-469O16.1		RP1-55C23.4		RP1-55C23.5	RP1-
55C23.7	RP1-6P5.2	RP1-78N10.2	RP1-86D1.2	RP1-86D1.3	RP1-86D1.4	RP1-86D1.5	RP1-8B1.4	RP3-323K23.3
RP3-330M21.5	RP3-337H4.6		RP3-351K20.3		RP3-351K20.4		RP3-403A15.1	RP3-
416F10.1	RP3-447E21.3	RP3-462C17.1	RP3-475N16.1		RP3-523C21.1		RP3-523C21.2	RP5-973N23.4
RP5-988G15.1	RPL24P4	RPL7L1	RPS12	RSPO3	SCARNA15	SLC22A7	SLC30A8	SNORA33
SNORA8	SNORD100	SNORD101	snoU13	SFR	STX7	TAAR1	TAAR2	TAAR3
TAAR5	TAAR6	TAAR7P	TAAR8	TAAR9	TAF8	TFEB	THEMIS	TJAP1
TRMT11	TTBK1	U4	U6	U7	USP49	VNN1	VNN2	VNN3
	Y_RNA	YIPF3	ZNF318					XPO5

PANTHER analysis: 83 mapped ids are found, 98 mapped ids are not found.

5.8.9.3 Genomic Gains of Pure Triple Negative DCIS and Triple Negative Invasive Breast Disease

There is overlap in the genomic gains present in TN pure DCIS (n=2/9) and TN Invasive breast disease (n=6/7).

1. p-values < 0.05;
2. A gene list is mapped from 11 genomic regions found on chromosomes 1, 6, 8, 10;
3. These regions encompass 133 genes altered in TN pure DCIS and TN invasive breast disease;

Genes:

5S_rRNA	7SK	AC099805.1	AL139392.1	AL359983.1	AP001207.1	C1orf150	CAPN11	CFLP4	CHML
EXO1	EYA1	FH	FMN2	GREM2	GRHL2	HSD17B7P1	KIF26B	KMO	MAP1LC3C
MRPL14	NCALD	NLRP3	OPN3	OR11L1	OR13G1	OR14A16	OR14A2	OR14C36	OR14K1
OR14L1P	OR1C1	OR2AJ1	OR2AK2	OR2B11	OR2C3	OR2G2	OR2G3	OR2G6	OR2L13
OR2L2	OR2L3	OR2L5	OR2L8	OR2M1P	OR2M2	OR2M3	OR2M4	OR2M5	OR2M7
OR2T10	OR2T11	OR2T12	OR2T2	OR2T29	OR2T3	OR2T33	OR2T34	OR2T35	OR2T4
OR2T6	OR2T8	OR2W3	OR2W5	OR6F1	OR9H1P	PLD5	RGS7	RP11-132G10.2	
RP11-177F11.1		RP11-314P12.2		RP11-323D18.1		RP11-323D18.4		RP11-323D18.5	
RP11-331N16.1		RP11-397A15.4		RP11-407H12.4		RP11-407H12.8		RP11-430I15.2	
RP11-430I15.4		RP11-435F13.1		RP11-435F13.2		RP11-438F14.1		RP11-438F14.3	

RP11-438H8.3	RP11-438H8.8	RP11-438H8.9	RP11-439E19.3	RP11-439E19.5
RP11-439E19.6	RP11-439E19.8	RP11-463J7.1	RP11-463J7.2	RP11-463J7.3
RP11-467I20.2	RP11-467I20.3	RP11-467I20.4	RP11-467I20.5	RP11-467I20.6
RP11-488L18.8	RP11-513D4.1	RP11-513D4.2	RP11-513D4.3	RP11-553N16.1
RP11-561I11.2	RP11-561I11.3	RP11-634B7.1	RP11-634B7.4	RP11-634B7.5
RP11-80B9.1	RP11-80B9.2	RP11-80B9.4	RP11-80B9.5	RP11-978I15.10
RP11-978I15.9	RSL24D1P4	SCCPDH	SLC29A1	SNORD112
	snoU13	TMEM63B	TRIM58	U5
				U6
VN1R17P	VN1R5	WDR64	XXyac-YR14BB7.1	Y_RNA

PANTHER analysis: 62 mapped ids are found, 71 mapped ids are not found.

There are genomic gains present in some samples of TN invasive breast disease (n=2/9), none of which are observed in TN pure DCIS (n=0/9).

1. p-values < 0.005;
2. A gene list is mapped from 2 genomic regions on chromosome 3;
3. These regions encompass 17 genes altered in TN invasive breast disease;

Genes:

AC133435.2	AC112504.2	RNF7	GRK7	ATP1B3	TFDP2	AC128648.1	GK5	RP11-340E6.1	RP11-340E6.2
RP11-144C9.1	RP11-271K21.2	RP11-271K21.4	RP11-271K21.7	RP11-343B5.1	AC133435.1	U6			

PANTHER analysis: Number of mapped IDs found 5, 12 mapped ids are not found.

Genomic Sc Gains in TN pure DCIS and TN Invasive Breast Disease

There are no Sc gains in TN pure DCIS in this series (n=0/9). Sc gains are, however, present in TN invasive breast disease (n=4/7).

1. p-values < 0.05;
2. A gene list is mapped from 14 genomic regions found on chromosomes 1, 3, 8;
3. These regions encompass 143 genes altered in TN invasive breast disease;

Genes:

5S_rRNA	AC023165.1	AC092898.1	AC092902.1	AC099805.1	AC112484.1	AC112504.2	AC117401.1	AC117508.1	AC128648.1
AC133435.1	AC133435.2	ACAD9	ADCY5	ALDH1L1	ALG1L	ATP1B3	C3orf15	C3orf22	C3orf37
CCDC14	CCDC48	CHCHD6	CHST13	CNBP	COPG	COX17	ENO1P3	EYA1	GK5
									GP9
GRAMD1C	GRK7	GS1-388B5.1	GS1-388B5.2	GS1-388B5.3	GS1-388B5.4	GS1-388B5.5			
GS1-388B5.6	GS1-388B5.7	GS1-388B5.8	GSK3B	H1FX	HEG1	hsa-mir-548i-1	ISY1		
ITGB5	KALRN	KIAA1257	KIAA1407	MUC13	MYLK	NR1I2	OSBPL11	PLXNA1	POPCD2
									PTPLB
QTRTD1	RAB43	RAB7A	RNF7	ROPN1	ROPN1B	RP11-124N2.1	RP11-124N2.2	RP11-124N2.3	
RP11-144C9.1	RP11-158I23.1	RP11-169N13.1	RP11-169N13.4	RP11-18H7.1	RP11-190C22.1				

RP11-197K3.1	RP11-202D20.1	RP11-221E20.3	RP11-221E20.4	RP11-221E20.5	RP11-255E6.1
RP11-271K21.2	RP11-271K21.4	RP11-271K21.7	RP11-290K4.2	RP11-340E6.1	
RP11-340E6.2	RP11-343B5.1	RP11-379B18.1	RP11-379B18.2	RP11-379B18.3	
RP11-379B18.4	RP11-379B18.5	RP11-390G14.1	RP11-434H6.2	RP11-435F17.1	
RP11-435F17.3	RP11-463J7.1	RP11-463J7.2	RP11-463J7.3	RP11-521J5.1	
RP11-529F4.1	RP11-605F14.1	RP11-605F14.2	RP11-605F14.3	RP11-666A20.1	RP11-666A20.3
RP11-666A20.4	RP11-689D3.4	RP11-71H17.1	RP11-71H17.6	RP11-722C17.1	
RP11-723O4.2	RP11-723O4.3	RP11-723O4.6	RP11-723O4.7	RP11-723O4.8	
RP11-767L7.1	RP11-767L7.2	RP11-775J23.1	RP11-775J23.2	RP11-9N20.1	RP13-685P2.4
RP13-685P2.5	RPS15AP16	RPS27P12	RSRC1	SHOX2	SLC12A8 SLC41A3
snou13	SNX4	TFDP2	TXNRD3	TXNRD3IT1	U1 U6 UMPS UROC1 Y_RNA
ZBTB20	ZDHHC23	ZNF148	ZXDC		

PANTHER analysis: 51 mapped ids are found, 92 mapped ids are not found.

5.8.9.4 Genomic Losses in TN pure DCIS and TN Invasive Breast Disease

There are genomic losses found in TN pure DCIS (n=1/9) also present in TN invasive breast disease (n=5/7).

1. p-values >0.05;
2. A gene list is mapped from 28 genomic regions found on chromosomes 4, 14, 15, 19;
3. These regions encompass 165 genes altered in TN pure DCIS and TN invasive breast disease;

Genes:

Y_RNA	5S_rRNA	7SK	ABLIM2	AC011912.1	AC012379.1	AC012379.2	AC013452.1	AC013452.2	AC018926.1
AC084882.1	AC090427.1	AC093680.1	AC093827.1	AC093848.1	AC097381.1	AC104059.1	AC104650.2	AC106868.1	AC107394.1
AC108516.1	AC109351.1	AC110766.1	ACOX3	ADH1A	ADH1B	ADH1C	ADH7	AF146191.4	AF250324.2
AFAP1	AFF1	AGPAT9	AIMP1	AL133444.1	AL160471.2	AL160471.3	ANXA3	AP001962.3	ARD1B
ARHGAP24	ARHGEF38	C15orf33	C4orf12	C4orf17	C4orf23	C4orf36	C4orf6	CCPG1	CEP152
CGRRF1	COPS2	CPZ	CRMP1	DAPP1	DNAJB14	DTWD1	EMR2	EVC	EVC2
FAM175A	FAM190A	FBN1	FGF5	FGF7	FOXA1	FRAS1	FRG1	FRMD6	GALK2
GCH1	GK2	GMPSP	GPR78	GSTCD	H2AFZ	hsa-mir-95	HSD17B11	HSD17B13	HTRA3
INTS12	KLHL8	MAPK10	MAPKSP1	MIPOL1	MTTP	NID2	NPNT	NUDT9	OR7E94P
PCDH7	PIGB	PRDM8	PTPN13	RAB27A	RASGEF1B	RG9MTD2	RP11-112L18.1		RP11-11N6.1
RP11-1258F18.1		RP11-15B17.1		RP11-15B17.4		RP11-162K6.1		RP11-174E22.1	
RP11-174E22.2		RP11-218C23.1		RP11-234K19.1		RP11-234K19.2		RP11-274J2.1	
RP11-311D14.1		RP11-311D14.2		RP11-377G16.2		RP11-390C19.1		RP11-397E7.1	
RP11-397E7.2		RP11-397E7.3		RP11-397E7.4		RP11-417M17.1		RP11-42A4.1	
RP11-438E5.1		RP11-452C8.1		RP11-45L9.1	RP11-463J17.1		RP11-476C8.1		RP11-476C8.2
RP11-476C8.3		RP11-529H2.1		RP11-610O8.1		RP11-617I14.1		RP11-656C2.1	
RP11-689K5.3		RP11-689P11.2		RP11-689P11.3		RP11-696N14.1		RP11-696N14.3	
RP11-710E1.1		RP11-710F7.2		RP11-710F7.3		RP11-722P15.1		RP11-766F14.1	

RP11-766F14.2	RP11-767N15.1	RP11-774O3.1	RP11-774O3.2	RP11-774O3.3
RP11-778J15.1	RP11-792D21.1	RP11-792D21.2	RP11-8L2.1	RP11-93M12.1
RP13-612N21.1				
RSL24D1	SAMD4A	SH3TC1	SHC4	SLC10A6
SNORA75	snoU13	SPARCL1	STK32B	TBCK
U5				
U6	U6atac	WDFY3		

PANTHER analysis: 67 mapped ids are found, 98 mapped ids are not found.

There are genomic losses found in TN invasive breast disease (5/7) not observed in TN pure DCIS (0/9).

1. p-values < 0.05;
2. A gene list is mapped from 8 genomic regions found on chromosomes 4, 8,14, 18, X;
3. These regions encompass 33 genes altered in TN invasive breast disease;

Genes:

AC018797.3	AF196969.5	AF196972.1	AF196972.9	AF213884.1	AF224669.3	AKAP6	AL592043.1	BMX	EBP
FAM47A	FTSJ1	MANBA	NFKB1	PIR	PORCN	RBM3	RP11-10L12.1		RP11-10L12.2
RP11-10L12.4		RP11-10L12.5		RP11-10L12.6		RP11-1148L6.5		RP11-1148L6.6	
RP11-281B1.2		RP11-305F18.1		RP11-545D19.1		RP13-202B6.2		SMAD4	snoU13
TBC1D25	UBE2D3	WDR13							

PANTHER analysis: 14 mapped ids are found, 19 mapped ids are not found.

5.8.9.5 Total Loss in TN pure DCIS and TN Invasive Breast Disease

There are total genomic losses present in TN pure DCIS (n=8/9) not observed in TN invasive breast disease (n=0/7).

1. p-values < 0.05;
2. A gene list is mapped from 23 genomic regions found on chromosomes 6, 8, 11, 13, 14, 17, 19;
3. These regions encompass 69 genes altered in TN pure DCIS;

Genes:

AARSD1	ABCC4	AC004659.1	AC005513.1	AC008734.1	AC008734.2	AC087650.1	AC087650.8	AC090427.1	AC100793.2
ACTL9	ADAM5P	ADAMTS10	AL136359.1	AL137001.1	AOC2	AOC3	ARL4D	BRCA1	
CACNA1A	CCDC105	CTD-2557P19.1		CTD-3199J23.2	EMR2	EMR3	FBXO33		
G6PC	GPC5	hsa-mir-2117IF135		MBD3L1	MUC16	MYO1F	OLFM4	OR2Z1	

OR4C3	OR4C45	OR4C5	OR4S1	OR4X1	OR4X2	OR7A10	OR7A11P	OR7A17
OR7A2P	OR7A5	OR7C1	OR7C2	PARK2	PSME3	RND2	RP11-24H2.1	RP11-24H2.2
RP11-301J16.2		RP11-301J16.3	RP11-301J16.5		RP11-301J16.7		RP11-442J17.4	
RPL27	RTN1	RUNCDC1	SLC1A6	SNORA40	U2	U6	VAT1	Y_RNA
ZNF333	ZNF558							

PANTHER analysis: 41 mapped ids are found, 28 mapped ids are not found.

There are total genomic losses present in TN invasive breast disease (n=7/7) not observed in TN pure DCIS (n=0/9).

1. p-values < 0.05;
2. A gene list is mapped from 30 genomic regions found on chromosomes 4, 8, 13, 14, 15, 17;
3. These regions encompass 112 genes altered in TN invasive breast disease;

Genes:

5S_rRNA	AC005324.1	AC005324.2	AC005324.4	AC005324.6	AC005324.7	AC007686.1	AC015922.1	AC015922.2	AC015922.7
AC090283.1	AC090283.3	AC096750.1	AC098487.1	ADCK1	AF111168.1	AF111168.3	AHSA1	AL049775.1	AL136160.1
AL357172.1	AL359180.1	ALKBH1	ARD1B	ATP5C1P1	C13orf23	C13orf36	C14orf148	C14orf156	C14orf174
C14orf178	C14orf184	CACNA1G	CATSPERB	CNOT6L	CTD-2175M1.1		DDIT4L	DYNLL1P1	FLRT2
GAPDHP34	GK2	GNG5P5	GSTZ1	hsa-mir-1260	IL6STP1	ISM2	KIAA1737	KRT222	KRT24
KRT25	LHFP	LMO7	MEIS3P1	MTUS1	NGB	NHLRC3	NRXN3	OR7E94P	POMT2
RFXAP	RP11-15B17.1	RP11-16L6.3	RP11-197L7.2		RP11-226P1.1		RP11-234K19.1		RP11-234K19.2
RP11-299L17.1		RP11-352E6.1		RP11-352E6.2	RP11-421P11.5		RP11-438E5.1		RP11-452C8.1
RP11-478K15.1		RP11-478K15.2		RP11-478K15.3		RP11-478K15.4		RP11-478K15.5	
RP11-499E18.1		RP11-527F15.1		RP11-588P8.1		RP11-610O8.1		RP11-624A4.1	
RP11-689K5.3		RP11-765K14.1		RP11-94C24.10		RP11-94C24.11		SLC39A8	SMAD9
SMARCE1	SNORA32	SNORA46	SNORA75	snoU13	SNW1	SPTLC2	STOML3	STON2	TBC1D26
TMEM63C	TRIM16	U3	U7	UBE2Z	VEGFC	VIPAR	Y_RNA	ZDHHC22	ZNF286
ZNF29P									

PANTHER analysis: 39 mapped ids are found, 72 mapped ids are not found.

5.8.9.6 CdLOH of TN pure DCIS and TN Invasive Breast Disease

There is CdLOH present in TN pure DCIS (n=1/9) also present in TN invasive breast disease (n=5/7).

1. p-values>0.05;
2. A gene list is mapped from 23 genomic regions found on chromosomes 4, 14, 15;

- These regions encompass 128 genes altered in TN pure DCIS and TN invasive breast disease;

Genes:

5S_rRNA	7SK	ABLM2	AC011912.1	AC012379.1	AC012379.2	AC013452.1	AC013452.2	AC018926.1	AC084882.1	
AC093680.1	AC093827.1	AC093848.1	AC097381.1	AC104059.1	AC106868.1	AC108516.1	AC109351.1	ACOX3	AF146191.4	
AF250324.2	AFAP1	AFF1	AGPAT9	AIMP1	AL133444.1	AL160471.2	AL160471.3	ANXA3	AP001962.3	
ARHGAP24	ARHGEF38	C15orf33	C4orf23	C4orf36	C4orf6	CCPG1	CEP152	CGRRF1	COPS2	CPZ
CRMP1	DAPP1	DNAJB14	DTWD1	EVC	EVC2	FBN1	FGF5	FGF7	FOXA1	
FRAS1	FRG1	FRMD6	GALK2	GCH1	GPR78	GSTCD	H2AFZ	HSD17B11	HSD17B13	
HTRA3	INTS12	KLHL8	MAPK10	MAPKSP1	MIPOL1	NID2	NPNT	NUDT9	PIGB	
PRDM8	PTPN13	RAB27A	RASGEF1B	RP11-112L18.1		RP11-15B17.1		RP11-15B17.4		
RP11-162K6.1		RP11-174E22.1		RP11-174E22.2		RP11-218C23.1		RP11-274J2.1		
RP11-311D14.1		RP11-311D14.2		RP11-377G16.2		RP11-390C19.1		RP11-397E7.1		
RP11-397E7.2		RP11-397E7.3		RP11-397E7.4		RP11-417M17.1		RP11-42A4.1		
RP11-45L9.1	RP11-463J17.1		RP11-476C8.1		RP11-476C8.2		RP11-476C8.3		RP11-529H2.1	
RP11-656C2.1		RP11-689K5.3		RP11-689P11.2		RP11-689P11.3		RP11-710E1.1		
RP11-710F7.2		RP11-710F7.3		RP11-766F14.2		RP11-767N15.1		RP11-774O3.1		
RP11-774O3.2		RP11-774O3.3		RP11-778J15.1		RP11-792D21.1		RP11-792D21.2		
RP11-8L2.1	RP13-612N21.1		RSL24D1	SAMD4A	SHC4	SLC10A6	snoU13	SPARCL1	STK32B	
TBCK	U5	U6	U6atac	Y_RNA+AA1:A128						

PANTHER analysis: 56 mapped ids are found, 72 mapped ids are not found.

There is CdLOH present in TN invasive breast disease (n=4/7) not observed in TN pure DCIS (0/9).

- p-values < 0.05;
- A gene list is mapped from 103 genomic regions found on chromosomes 4, 5, 8, 9, 10, 12, 13, 14, 15, 16, 18, X;
- These regions encompass 433 genes altered in TN invasive breast disease;

Genes:

5S_rRNA	AC003658.1	AC004053.1	AC006160.1	AC006160.2	AC006160.4	AC006160.5	AC006445.1	AC006445.6	AC006445.7	
AC006445.8	AC007126.1	AC007956.2	AC008245.1	AC010296.1	AC013558.1	AC018413.1	AC018558.1	AC019131.2	AC020698.1	
AC023824.2	AC034110.1	AC034110.2	AC068473.1	AC090360.1	AC091551.1	AC091602.1	AC091646.1	AC092621.1	AC092846.1	
AC093628.1	AC093848.1	AC096742.1	AC098830.1	AC105423.1	AC106785.1	AC106785.10		AC106785.14		
AC106785.18		AC106785.20		AC106785.22		AC106785.25		AC106785.27		
AC106785.28		AC106785.6	AC106868.1	AC111152.1	AC114688.1	AC116553.1	AC134978.1	AC139100.1	ACTBP1	
ACYP1	ADH1A	ADH1B	ADH1C	ADH4	ADH5	ADH7	ADNP2	AF196970.3	AF196972.3	
AF196972.4	AF213884.1	AF241726.2	AF241726.4	AF241726.6	AKAP6	AL096700.1	AL096700.2	AL121578.2	AL132716.1	
AL132716.2	AL132777.1	AL132777.2	AL132988.1	AL132988.2	AL132988.3	AL132988.4	AL359219.1	AL445487.1	AL591394.1	
AL929101.1	ANXA10	AP002075.1	ARAF	ASB5	ATP8A1	ATP9B	B3GALT1	BANK1	BCL2	
BDH2	BOD1L	C13orf26	C13orf33	C14orf149	C14orf38	C14orf53	C15orf43	C18orf22	C4orf37	

C8orf75	CDH7	CENPE	CENPVL1	CFP	CHST7	CISD2	CLRN2	COMMD10	CTB-118N6.3
CTC-339F2.2		CTD-2022H16.1		CTD-2522E6.4		CTDP1	CXorf24	CXorf25	CXorf31
CXXC4	CYP4V2	DCTN2	DDIT3	DDX60	DDX60L	DLST	DUOX1	DUOX2	DUOXA1
DUOXA2	DUSP4	EEF1DP3	EGLN3	EIF2B2	EIF4E	ELK1	EMCN	FAM149A	FAM164C
FAM177A2	FAM184B	FANCB	FCF1	FRY	FTLP16	FTSJ1	GALR1	GBA3	GLRA2
GPR125	GPR64	GSPT2	hsa-mir-1255a		hsa-mir-221	hsa-mir-222	hsa-mir-616	HSBP1L1	HSP90AB2P
HSPH1	INE1	JKAMP	KCNG2	KCNH5	KCNIP4	KIAA0317	KIF5A	KLF12	KLHL4
KLKB1	KRT8P14	LAP3	LDB2	LRRC9	LTBP2	MAGED1	MAGED4	MAGED4B	MAP3K15
MARS	MBD6	MBP	MED28	METAP1	MEX3C	MLH3	MNAT1	MOSPD2	MTNR1A
NDUFB11	NEK9	NFATC1	NFKB1	NHEDC1	NHEDC2	NKX3-2	NOVA1	NPAS3	NUS1P1 OTC
PARD6G	PCDH7	PCTK1	PDHA1	PGF	PHF16	PHKA2	PPARGC1A	PPEF1	PPP3CA
PQLC1	PROX2	QDPR	RAB28	RAD23B	RBM10	RGN	RP11-109E24.1		RP11-10G12.1
RP11-10G12.2		RP11-10L12.4		RP11-1148L6.6		RP11-114H20.1		RP11-120A1.1	
RP11-123O22.1		RP11-1299A16.1		RP11-1299A16.2		RP11-141E13.1		RP11-159J2.1	
RP11-159J2.2		RP11-162A12.1		RP11-167N19.1		RP11-167N19.2		RP11-173M11.1	
RP11-173M11.2		RP11-173P16.2		RP11-17E2.2		RP11-196I18.2		RP11-196I18.3	
RP11-196I18.4		RP11-198M15.1		RP11-200A24.1		RP11-207N4.2		RP11-207N4.3	
RP11-215A19.1		RP11-22B10.3		RP11-234P3.2		RP11-234P3.3		RP11-234P3.4	
RP11-234P3.5		RP11-245M24.1		RP11-24D11.1		RP11-252M21.1		RP11-252M21.3	
RP11-252M21.4		RP11-252M21.6		RP11-252M21.7		RP11-289C17.1		RP11-289C17.2	
RP11-292B1.1		RP11-292B1.2		RP11-297P16.1		RP11-297P16.3		RP11-297P16.4	
RP11-305F18.1		RP11-30C8.1		RP11-30C8.2		RP11-310I9.1		RP11-311P8.1	
RP11-311P8.2		RP11-315A17.1		RP11-324B6.1		RP11-328K4.1		RP11-333E5.1	
RP11-333E5.2		RP11-341G5.1		RP11-367C11.2		RP11-380P13.1		RP11-380P13.2	RP1-
138A5.1	RP11-38O23.3		RP11-38O23.4		RP11-38O23.5		RP11-38O23.7		RP11-417L14.1
RP11-428B4.2		RP11-433O3.1		RP11-451L9.2		RP11-451L9.3		RP11-45F23.1	
RP11-462G22.1		RP11-473D24.1		RP11-474L7.4		RP11-495L18.2		RP11-498M5.2	
RP11-499E18.1		RP11-49F1.1	RP11-508N12.2		RP11-508N12.3		RP11-508N12.4		RP11-552E4.2
RP11-552E4.3		RP11-552E4.4		RP11-556G22.1		RP11-556G22.2		RP11-556G22.3	
RP11-56H2.1		RP11-571E6.1		RP11-571E6.3		RP11-571E6.4		RP11-571L19.2	
RP11-571L19.4		RP11-576E20.1		RP11-617I14.1		RP11-619J20.1		RP11-644A7.1	
RP11-655M19.1		RP11-665C14.1		RP11-665C14.2		RP11-665I14.1		RP11-692P14.1	
RP11-696N14.1		RP11-696N14.3		RP11-6C14.1		RP11-703G6.1		RP11-729M20.1	
RP11-75A9.1		RP11-75A9.2		RP11-75A9.3		RP11-769N22.1		RP11-799A12.1	
RP11-799A12.2		RP11-87F15.2		RP1-212G6.4		RP1-30G7.2	RP13-43E11.1		RP13-479F17.2
RP13-497K6.1		RP13-928P6.3		RP1-54B20.7		RP1-71L16.1	RP1-71L16.2	RP1-71L16.3	RP1-71L16.6
RP3-393P12.1		RP3-393P12.2		RP3-393P12.3		RP3-436M11.3		RP3-499B10.3	RP3-
499B10.4	RP3-499B10.5		RP4-733D15.1		RP4-774G10.1		RP5-1158E12.1		RP5-1158E12.2
RP5-1158E12.3		RP5-972B16.2		RP6-227L5.1	RP6-227L5.2	RP6-99M1.1	RPGR	RPS6KL1	RTN1
RXFP2	SALL3	SEMA6A	SLC25A21	SLC38A5	SLC38A6	SLC7A1	SLC9A7	SMAD4	SNORA11
SNORA25	SNORA31	SNORA7	SNORA75	SNORD37	snosnR60_Z15		snoU13	SOC56	SORD
SPACA5	SPACA5B	SPCS3	SRPX	SSX5	SSX6	STAMBPL1	SUV39H1	SYN1	SYTL5
TACR3	TIMP1	TJP1	TLR3	TMED10	TRMT5	TSHZ1	SPAN5	SPAN7	TXNL4A U1
U12	U2	U3	U4atac	U6	U7	U70984.1	UBA1	UBE2D3	UNC13C
UPRT	USP11	UXT	WAS	WDR72	Y_RNA	YLPM1	Z98304.1	ZADH2	ZDHC15
ZNF157	ZNF182	ZNF236	ZNF41	ZNF462	ZNF516	ZNF630	ZNF673	ZNF674	ZNF81

PANTHER analysis: 163 mapped ids are found, 270 mapped ids are not found.

5.8.9.7 CnLOH of TN pure DCIS and TN Invasive Breast Disease

There is CnLOH present in TN pure DCIS (n=9/9) not observed in TN invasive breast disease (n=0/7).

1. p-values < 0.05;
2. A gene list is mapped from 110 genomic regions found on chromosomes 1, 2, 3, 6, 7, 10, 11, 12, 20, 21;
3. These regions encompass 891 genes altered in TN pure DCIS;

Genes:

5S_rRNA 7SK ABHD11 AC004975.1 AC004979.1 AC004979.2 AC004979.3 AC004979.4 AC004979.5 AC004979.6 AC004979.7
AC005537.2 AC005586.2 AC006021.1 AC007092.1 AC007238.1 AC007560.1 AC007680.2 AC007682.1 AC007682.2 AC008173.1
AC010087.1 AC010087.4 AC010087.5 AC012361.1 AC012368.1 AC012368.2 AC013251.1 AC013251.2 AC016727.3 AC016894.1
AC017083.1 AC017083.2 AC017083.3 AC018462.2 AC018462.3 AC022210.2 AC023115.1 AC023115.3 AC023165.1 AC025594.1
AC025594.2 AC064868.1 AC067950.1 AC069304.1 AC073422.1 AC073464.11 AC073464.9 AC073846.1 AC073846.2
AC073846.3 AC074289.1 AC078811.1 AC078941.1 AC079112.1 AC079807.4 AC083868.1 AC083883.1 AC092898.1 AC092902.1
AC092910.1 AC092910.2 AC092958.1 AC093087.1 AC093168.1 AC097463.1 AC097463.2 AC106874.1 AC106874.3 AC107081.1
AC107081.2 AC107081.5 AC108039.1 AC108039.2 AC108039.3 AC108462.1 AC108697.1 AC108729.1 AC108739.1 AC112484.1
AC112694.1 AC114752.1 AC114808.2 AC114808.3 AC117401.1 AC117472.1 AC117472.2 AC117508.1 AC126182.1 AC131571.1
AC131571.2 AC138655.4 AC144441.1 ACAD9 ACTG2 ACTGP9 ACTR3C ADCY5 AF131217.1 AF227510.7 AGK
AGTR1 AKR1A1 AL008729.1 AL022170.1 AL023583.1 AL023807.1 AL023807.2 AL024498.1 AL034372.1 AL035670.1
AL049635.1 AL049635.2 AL049710.1 AL049710.2 AL049710.4 AL050335.1 AL050335.2 AL109918.1 AL121946.1 AL133268.1
AL136230.1 AL136305.1 AL137003.1 AL138724.3 AL139044.2 AL157773.1 AL160037.2 AL160400.1 AL160400.2 AL162579.1
AL354680.1 AL357079.1 AL357079.2 AL357497.1 AL359316.1 AL445669.2 AL589723.1 AL590062.1 AL603888.1 AL604028.1
AL604028.2 AL645811.1 ALG1L ALK ANKRD20B APOB ARGFX ARTN ASB3 ASS1P1 ATP5LP2
ATP5LP5 ATP6V0B ATP6V1F ATXN1 B3GNT1 B4GALT2 BAK1 BBS9 BCL7B BMP5
BPESC1 BTF3L4 C1D C1orf210 C1orf50 C1orf84 C2orf43 C2orf89 C3orf16 C3orf22 C3orf33 C3orf37
C3orf72 C6orf114 C6orf142 C6orf227 C6orf52 C7orf29 C7orf55 CAP1 CAP2 CCDC136 CCDC14
CCDC17 CCDC23 CCDC24 CCDC30 CCDC48 CCDC90A CCT4 CD2AP CD83 CD96 CEP70
CHCHD6 CHST13 CLDN19 CLDN3 CLDN4 CLEC5A CMAH CNBP COL21A1 COL9A2 COMMD1
COMMD2 COPG COX6B1P1 CTD-2021J15.1 CTD-2021J15.2 CYP26B1 CYP39A1 DEFB110
DEFB112 DEFB113 DEFB114 DEFB133 DEK DNAH6 DNAJB6 DNAJC30 DPH2 EBNA1BP2
EDN1 EEF1A1P25 EFHC1 ELOVL2 ELP4 ENO1P3 EPCAM ERBB3 ERMAP ESYT3 FABP7
FAIM FAM161A FAM65B FAM8A1 FBXO40 FLNC FOXL2 FOXN2 FSTL1 GAPDHP39
GCLC GCM2 GCNT2 GCNT6 GFOD1 GFRAL GGNBP1 GIMAP2 GIMAP4 GIMAP6 GIMAP7
GIMAP8 GJA9 GMNN GMPR GMP5 GP9 GPBP1L1 GPR110 GPR111 GPR115
GPR116 GPR156 GRM8 GS1-388B5.1 GS1-388B5.2 GS1-388B5.3 GS1-388B5.4
GS1-388B5.5 GS1-388B5.6 GS1-388B5.7 GS1-388B5.8 GSK3B GSTA1 GSTA2
GSTA3 GSTA5 GSTA6P GTF2E1 GTF2H1 H1FX HCRT2 HEG1 HGD HIPK2
HIVEP1 HLA-DQA1 HLA-DQB1 HMGCLL1 HMGCS2 hsa-mir-133b hsa-mir-206 hsa-mir-548a-1 hsa-mir-548i-1
hsa-mir-559 hsa-mir-761 HSPDP7 HY1 IGSF22 IL17A IL17F IMMP1L IPO13 IPP IQCJ IRF5
ISY1 ITGB5 ITPR3 JARID2 JHDM1D KALRN KCNK12 KCP KDM1B KDM4A KIAA0319L
KIAA0467 KIAA1147 KIAA1257 KIF13A KLHL31 KLRAQ1 KLRG2 KT112 LDHA LDHAL6A
LDHC LEPRE1 LRFN2 LRRC1 LRRC16A LRRC4C LRRC58 LRRC61 LUC7L2 MAK
MAST2 MCM3 MED8 MEP1A MFSD2A MGAM MKRN1 MLXIPL MME MRAS MRPL32

	MRPS22	MSH2	MTCO3P1	MUC13	MUTYH MYCBP	MYL6P4	MYLIP	MYLK	N6AMT1	
	NAP1L1P3	NASP	NCDN	NCRNA00240	NEDD9	NHLRC1	NOL7	NRD1	NRXN1	NUP153 OFCC1
OR5AK2	OR9A1P	OR9A3P	OR9A4	OSBPL11	OSBPL9	P2RY1	PABPC1P10	PAK1IP1	PAQR8	
	PARP12	PAX6	PDIA4	PDIA5	PELI1	PFN2	PHACTR1	PHGDH	PIK3CB	PIK3R3
PLA2G7	PLCH1	PLCXD2	PLXNA1	PNO1	PPCS	PPIH	PPP3R1	PPT1	PRIM2	PRR23A
	PRR23B	PRR23C	PRSS37	PSMB2	PTPLB	PTPN5	PTPRF	RAB19	RAB3B	RAB43
	RAB7A	RABL3	RANBP9	RARRES1	RARRES2	RBM24	RCAN2	REPIN1	RHBDL2	RIMKLA
	RNF13	RNF144B	RNF160	RNF182	ROPN1	RP11-100F22.1	RP11-10Q22.2	RP11-115D7.3		RLF
	RP11-117F22.1	RP11-117L15.1	RP11-117L15.3	RP11-121P10.1	RP11-1220K2.2	RP11-125M16.1				
	RP11-127P7.2	RP11-128A6.2	RP11-128A6.3	RP11-138N21.1	RP11-146I2.1	RP11-146I2.2	RP11-14H3.3			
	RP11-157D18.2	RP11-163G10.2	RP11-163G10.3	RP11-163G10.4	RP11-167H9.3	RP11-167H9.4				
	RP11-174O3.1	RP11-174O3.3	RP11-174O3.4	RP11-184I16.2	RP11-184I16.3	RP11-184I16.4				
	RP11-18H7.1	RP11-191A15.1	RP11-191A15.2	RP11-191A15.3	RP11-191A15.4	RP11-192K2.3				
	RP11-197K3.1	RP11-202D20.1	RP11-219E24.1	RP11-221E20.3	RP11-221E20.4	RP11-221E20.5				
	RP11-228O6.2	RP11-232M24.1	RP11-255N4.2	RP11-255N4.3	RP11-268F1.2					
	RP11-268F1.3	RP11-280H21.1	RP11-282K6.3	RP11-290K4.1	RP11-290K4.2	RP11-291L19.1				
	RP11-306L14.1	RP11-309L24.2	RP11-309L24.4	RP11-309L24.6	RP11-310F12.1					
	RP11-311B18.1	RP11-321G3.1	RP11-322N21.2	RP11-330A16.1	RP11-33K5.1					
	RP11-342M1.2	RP11-342M1.3	RP11-342M1.4	RP11-345L23.1	RP11-359H3.1	RP11-359H3.4				
RP11-359N11.1	RP11-359N11.2	RP11-360O19.1	RP11-360O19.4	RP11-360O19.5	RP11-365F18.1					
	RP11-365F18.3	RP11-367G6.1	RP11-379B18.1	RP11-379B18.2	RP11-379B18.3					
	RP11-379B18.4	RP11-379B18.5	RP11-383F6.1	RP11-384F7.1	RP11-384F7.2					
	RP11-385F7.2	RP11-38P22.2	RP11-390G14.1	RP11-392A23.4	RP11-392A23.5	RP11-397G17.1				
RP11-13D10.2	RP1-13D10.3	RP1-13D10.4	RP1-13D10.5	RP11-401E14.1	RP11-401E14.2					
	RP11-402J7.2	RP11-402J7.3	RP11-40M23.1	RP11-411K7.1	RP11-411K7.2	RP11-411K7.4				
	RP11-411K7.5	RP11-418B12.1	RP11-428L9.1	RP1-142O9.1	RP1-142O9.2					
	RP11-434H6.2	RP11-446F17.3	RP11-446F3.2	RP11-451G4.2	RP11-451G4.3	RP11-466L17.1				
	RP11-470H9.1	RP11-500K19.1	RP11-501I19.3	RP11-501I19.4	RP11-501O2.1					
	RP11-501O2.2	RP11-501O2.3	RP11-501O2.4	RP11-501O2.5	RP11-505J9.1					
	RP11-506B15.4	RP11-506B15.6	RP11-511P7.2	RP1-151F17.1	RP11-521J5.1	RP11-526P5.1				
	RP11-526P5.2	RP11-529F4.1	RP11-529G21.4	RP1-152L7.1	RP11-152L7.3	RP11-152L7.5	RP11-152L7.6	RP11-543F8.1		
	RP11-543F8.2	RP11-548O1.1	RP11-548O1.3	RP11-550C4.6	RP11-560J1.1	RP11-567P17.1				
	RP11-567P17.2	RP11-585F20.1	RP11-598F17.1	RP11-605F14.1	RP11-605F14.2					
	RP11-605F14.3	RP11-622P13.2	RP11-630I5.1	RP11-632K21.1	RP11-632K21.3					
	RP11-635I10.1	RP11-637O19.2	RP11-639B1.1	RP11-639F1.1	RP11-639F1.2	RP11-639F1.3				
RP11-63E16.1	RP11-649A16.1	RP11-649A16.5	RP11-651P23.2	RP11-651P23.3	RP11-651P23.4					
	RP11-651P23.5	RP11-651P23.6	RP11-666A20.1	RP11-666A20.3	RP11-666A20.4					
	RP11-674E16.1	RP11-681L4.1	RP11-689D3.4	RP11-69C17.3	RP11-69C17.4					
	RP11-71H17.1	RP11-71H17.6	RP11-71N10.1	RP11-722C17.1	RP11-723C11.2					
	RP11-723O4.2	RP11-723O4.3	RP11-723O4.6	RP11-723O4.7	RP11-723O4.8					
	RP11-735A5.1	RP11-767N6.2	RP11-767N6.7	RP11-771D21.2	RP11-775J23.1					
	RP11-775J23.2	RP11-779P15.1	RP11-779P15.2	RP11-785H2.1	RP11-788A4.1					
	RP11-788A4.2	RP11-793K1.1	RP11-795J1.1	RP11-795J1.2	RP11-79N23.1	RP11-7O11.3				
	RP1-180E22.3	RP11-812I20.2	RP11-90C4.1	RP11-90C4.2	RP11-90C4.3	RP11-90C4.4				
	RP1-190J20.2	RP11-91A18.1	RP11-91A18.4	RP11-91A18.5	RP11-9N20.1	RP1-228H13.1				
	RP1-228H13.2	RP1-245M18.2	RP1-257A7.4	RP1-273P12.1	RP1-273P12.3	RP1-27K12.2				
	RP1-27K12.4	RP1-290I10.4	RP1-290I10.5	RP1-290I10.7	RP1-304B14.3	RP13-476E20.1				
RP13-685P2.4	RP13-685P2.5	RP1-39G22.4	RP1-39G22.5	RP1-62D2.3	RP1-71H19.2	RP1-				
92O14.6	RP3-335N17.2	RP3-342P20.2	RP3-347E1.2	RP3-365O12.2	RP3-380E11.2	RP3-				
425P12.1	RP3-425P12.2	RP3-425P12.4	RP3-425P12.5	RP3-437C15.1	RP3-448I9.1	RP3-448I9.2	RP3-			

PANTHER analysis: 335 mapped ids are found, 497 mapped ids are not found.

1. p-values < 0.05;
2. A gene list is mapped from 30 genomic regions found on chromosomes 1, 4, 7, 10, 14, 18, 21;
3. These regions encompass 379 genes altered in TN invasive breast disease;

5S_rRNA 7SK	ABCA4	ABCD3	ABLIM1	AC005071.1	AC005071.2	AC005071.3	AC005879.1	AC006557.1	AC007922.2	AC007956.2
	AC016839.1	AC018797.3	AC022884.1	AC023983.1	AC069281.1	AC069281.2	AC073585.1	AC073588.1	AC073842.18	AC073842.19
	AC090363.1	AC091588.1	AC093577.3	AC093577.5	AC093577.7	AC093577.8	AC096947.1	AC104961.1	AC105247.1	ACADSB ACTL6B
AFAP1L2	AFG3L2	AGFG2	AL109761.3	AL109761.5	AL117351.1	AL117351.2	AL117351.3	AL117351.4	AL117351.5	
	AL117351.6	AL117351.7	AL133384.1	AL157788.1	AL158014.1	AL353664.1	AL355598.1	AL355861.1	AL359747.1	AL359836.1
	AL603764.1	AL731543.1	ALG14	AMAC1L1	ANKRD62	AP000431.1	AP000431.2	AP000432.2	AP000457.1	AP000568.2
	AP000745.1	AP000855.4	AP000946.2	AP000963.2	AP000963.3	AP000998.1	AP000998.2	AP001029.1	AP001120.1	AP001171.1
	AP001439.2	AP001441.1	AP001442.2	AP001525.1	AP001525.2	AP001604.3	AP001605.4	AP001605.6	AP001607.1	AP002414.1
	AP002414.2	AP005264.2	AP4M1	APP	ARHGAP29	ARMS2 ATE1	BAG3	BCAR3	BTBD16	
	BTG3	BUB3	C10orf119	C10orf120	C10orf46	C10orf82	C10orf84	C10orf88	C10orf96	C18orf1 C18orf19
C18orf45	C1orf87	C1QBPP	C21orf131	C21orf37	C21orf91	C7orf43	C7orf47	C7orf51	C7orf59	C7orf61
	CABLES1	CACNA2D1	CASC2	CCDC18	CDH2	CELF4	CHODL	CHST9	CIDEA	CNN3 CNPY4
	COPS6	CTA-109P11.1	CTAGE1	CTB-161A2.2			CTB-161A2.3		CTB-161A2.4	
	CUZD1	CXADR	CYP2J2	DLST	DMBT1 DNMBP	DNTTIP2	DR1	EIF3A	EIF4A1P	
	EMX2	EMX2OS	ENO4	F3 FAM24A	FAM24B	FAM45A	FAM69A	FBXQ24	FCF1	FDPSP FGFR2

FGGY	FNBP1L	GAL3ST4	GAPDHP29	GATA6	GATS	GCLMGFRA1	GIGYF1	GNAL	GNB2	
GPC2	GRK5	HMX2	HMX3	HOOK1	HRH4	hsa-mir-106b	hsa-mir-1-2	hsa-mir-133a-1	hsa-	
mir-25	hsa-mir-320c-2	hsa-mir-760	hsa-mir-93	HSPA12A	HTR7	HTRA1	IKZF5	IMPA2	IMPACT	INADL
INPP5F	KCNK18	KCTD1	KIAA0317	KIAA1598	KRT18P2	LRCH4	LTBP2	MANBA	MBLAC1	
MC2R	MC5R	MCM7	MEPCE	MIB1	MIR1255B1	MOSPD3	MPPE1	MTF2	NANOS1	NCAM2
NCRNA00093	NCRNA00113	NCRNA00157	NFIA	NFKB1	NSMCE4A	ORC5L	OSBPL1A		PCOLCE	
PDZD8	PILRA	PILRB	PLEKHA1	PMS2L1	PNLIP	PNLIPRP1	PNLIPRP2	PNLIPRP3	PPIAP	
PRDX3	PRLHR	PROX2	PRSS7	PSMA8	PSTK	PVRIG	RAB11FIP2	RELN	RGS10	
RNMT	RP11-107C16.2	RP11-10L12.1		RP11-10L12.2		RP11-10L12.4		RP11-10L12.5		
	RP11-10L12.6	RP11-145M4.2	RP11-145M4.3		RP11-148B18.1		RP11-148B18.3		RP11-148B18.4	
RP11-162A23.5	RP11-179H18.2		RP11-179H18.4		RP11-179H18.5		RP11-179H18.7		RP11-179H18.8	
	RP11-198M6.2	RP11-198M6.5		RP11-215A21.2		RP11-245J24.1		RP11-268I9.1		
	RP11-26N15.1	RP11-282I1.1	RP11-295O23.2		RP11-295O23.4		RP11-313A24.1		RP11-317F20.2	
RP11-319I23.2	RP11-319I23.3		RP11-328K15.1		RP11-33D13.2		RP11-354M20.3		RP11-366L18.1	
	RP11-36N22.1	RP11-36N22.3		RP11-430G17.1		RP11-436K8.1		RP11-44M6.1		
	RP11-44M6.3	RP11-465K1.2	RP11-470E16.1		RP11-488P3.1		RP11-498B4.2		RP11-498B4.5	
	RP11-498B4.7	RP11-498J9.1		RP11-498J9.2		RP11-498J9.4		RP11-500G22.2		
	RP11-501J20.2	RP11-501J20.3		RP11-501J20.5		RP11-506M12.1		RP11-539I5.1		
	RP11-539I5.2	RP11-564D11.3		RP11-567J24.4		RP11-575B7.2	RP11-575B7.3		RP11-57H12.2	
	RP11-5G18.2	RP11-62L18.3		RP11-681N23.1	RP11-758P17.2		RP11-758P17.3		RP11-781P14.3	
	RP11-786F14.1	RP11-78A18.2	RP11-78O9.1		RP11-79M19.2		RP11-86H7.1		RP11-86H7.2	
	RP11-86H7.3	RP11-86H7.6		RP11-86H7.7		RP13-530H6.2		RP4-612C19.1		RP4-
612C19.2	RP4-639F20.1		RP4-639F20.3		RP4-668G5.1		RP4-713B5.2		RP4-717I23.2	
RP4-717I23.3	RP4-782L23.1		RP4-782L23.2		RP5-1033H22.2		RP5-833A20.1	RP5-837O21.1		RP5-
837O21.2	RP5-837O21.3		RP5-837O21.4		RP5-837O21.5		RP5-837O21.6		RP5-837O21.7	
	RP5-837O21.8	RP5-837O21.9		RPL10P1	RPL37P3	RPL37P4	RPS15AP5	RPS3AP1		RPS6KL1
SEC23IP	SFXN4	SLC18A2	SLC44A3	SLC6A6P	SLMO1	SNORA19	SNORA7	SNORA73	snoU13	
SPDYE3	SPIRE1	SS18	STAG3	TACC2	TAF4B	TAF6	TFR2	TGFBR3	TIAL1	TM2D1
TMEM56	TSC22D4	TUBB6	U1	U3	U6	U6atac	U7	UBE2D3	VAX1	Y_RNA
YLPM1	ZCWPW1	ZNF3	ZNF519	ZNF521	ZSCAN21					

PANTHER analysis: 152 mapped ids are found, 227 mapped ids are not found.

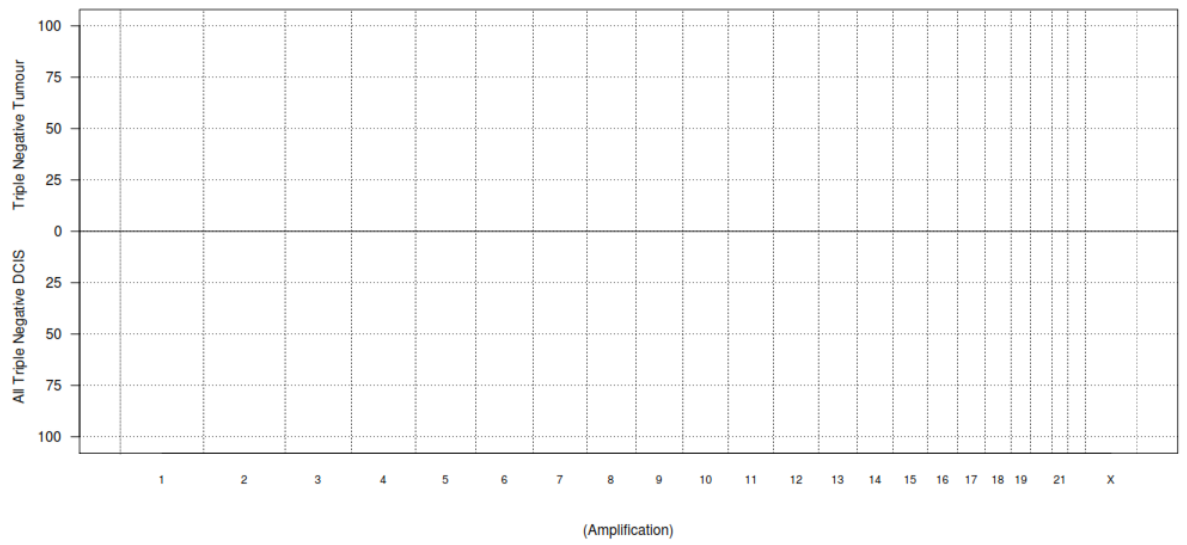
5.8.10 Copy Number Aberrations in All Triple Negative DCIS compared to Triple Negative Invasive breast disease

This series examines the difference between TN DCIS and TN invasive breast disease.

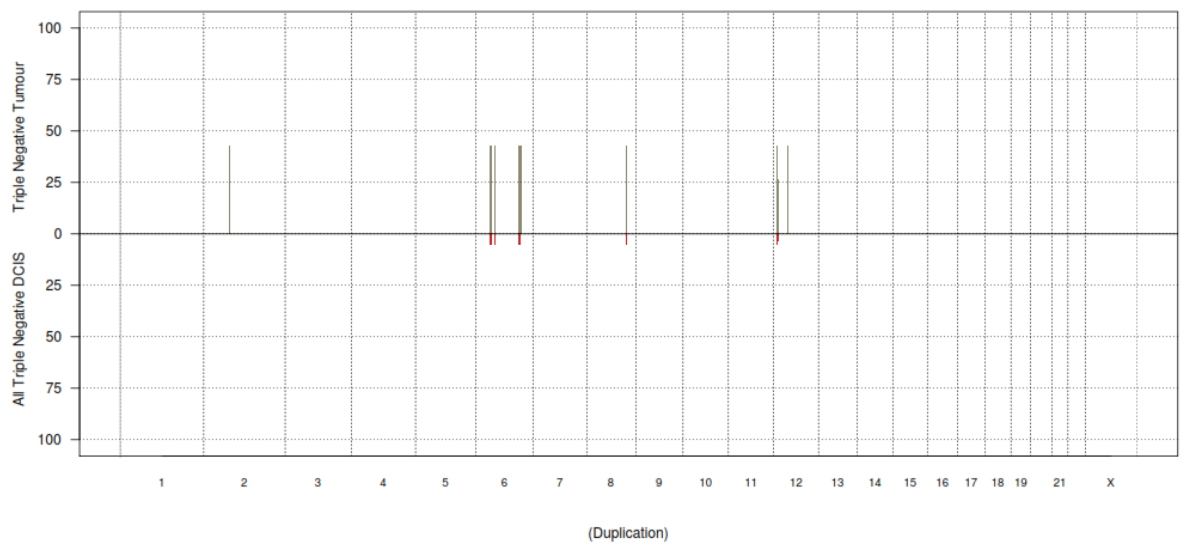
Frequency plots showing copy number aberrations between TN DCIS and TN invasive breast disease were provided by King's College London Bioinformatics department; Figure 43.

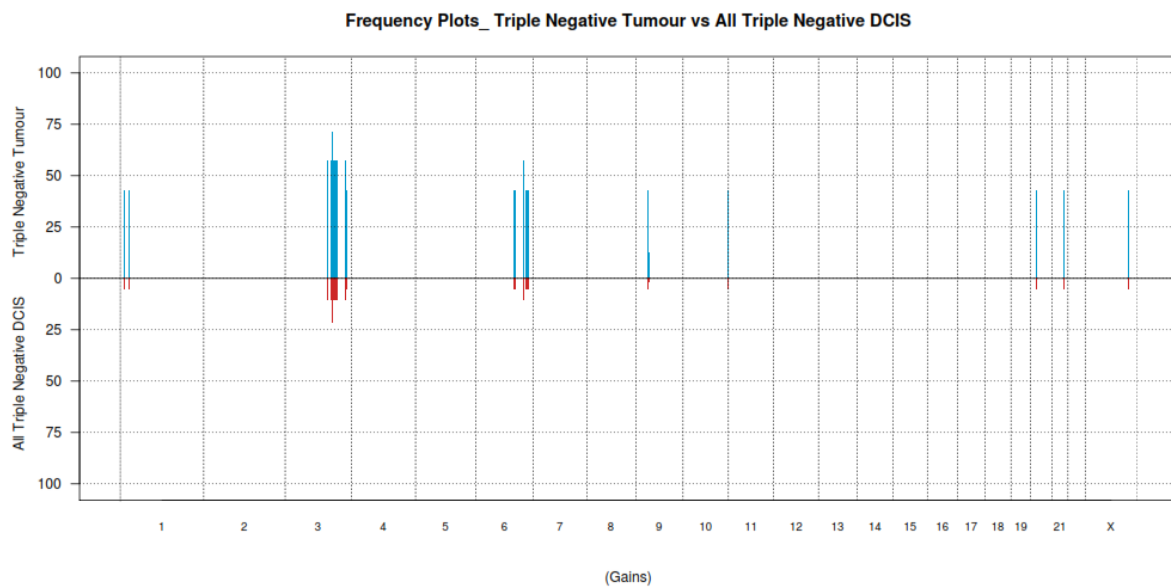
Figure 43: Frequency plots showing copy number aberrations between triple negative DCIS and triple negative invasive disease (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 264-266).

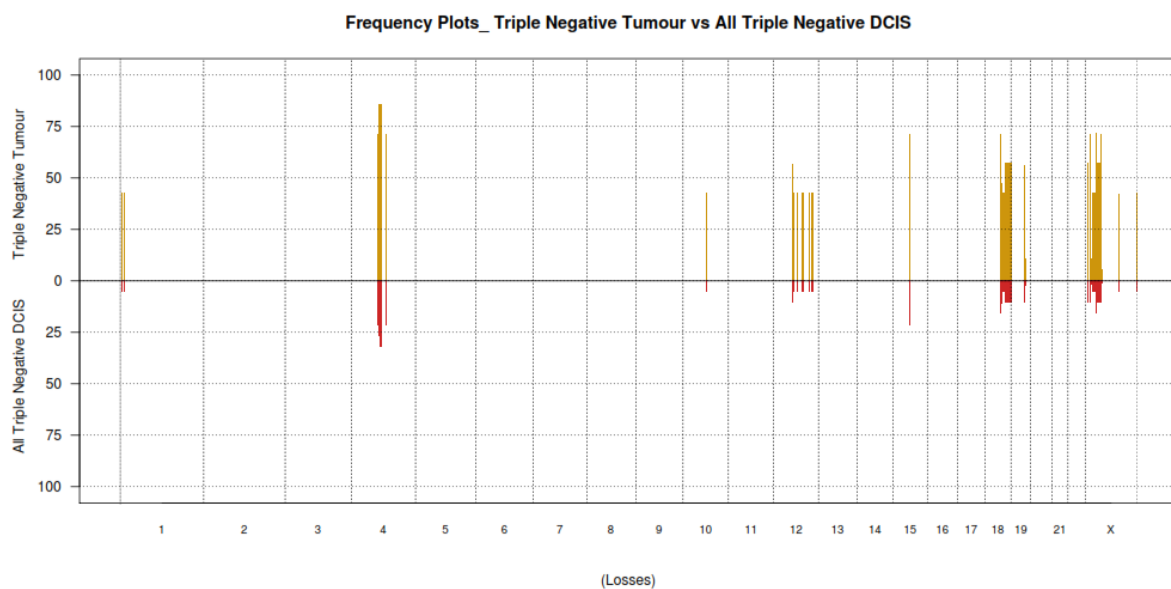
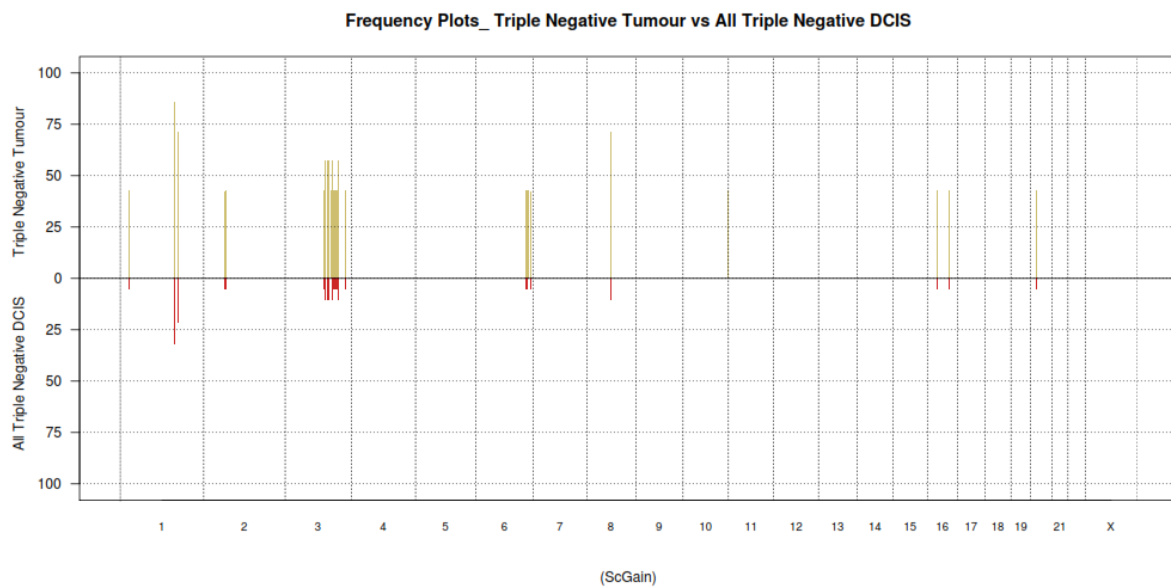
Frequency Plots_ Triple Negative Tumour vs All Triple Negative DCIS

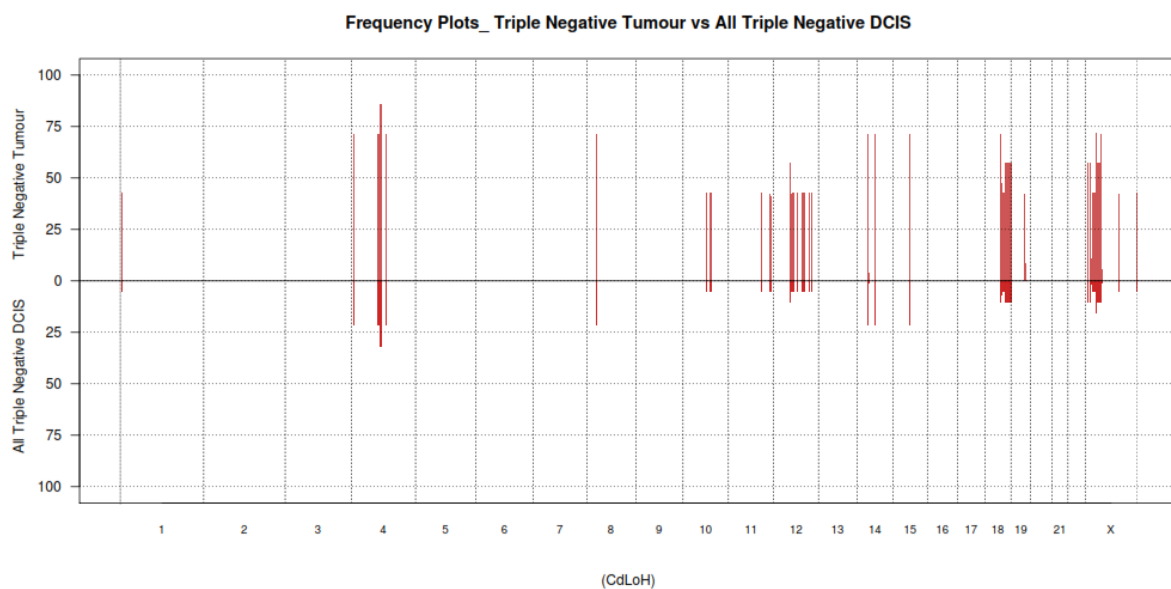
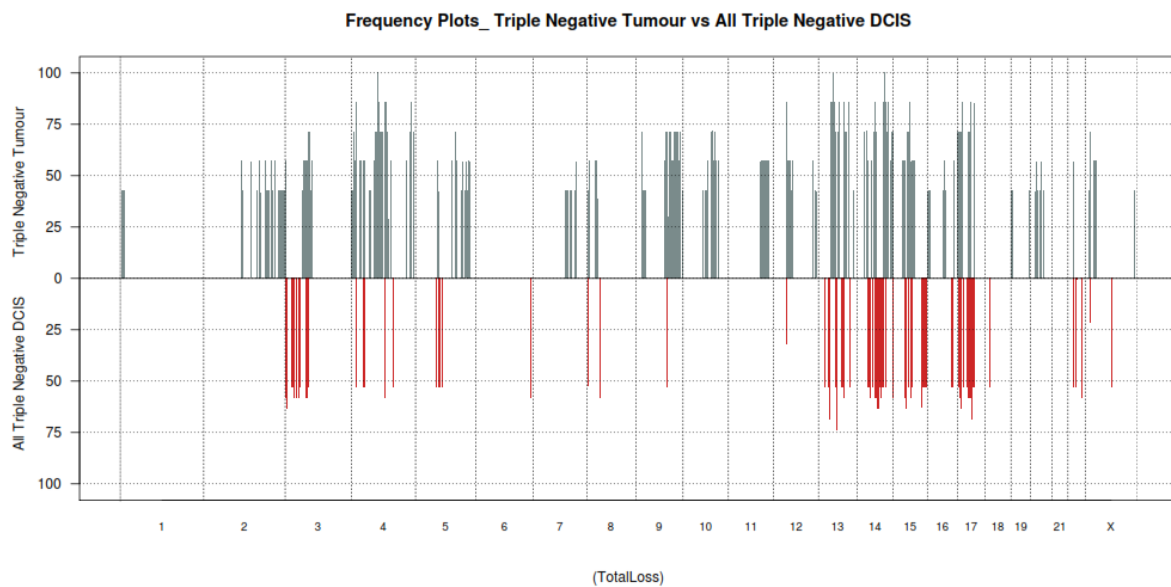


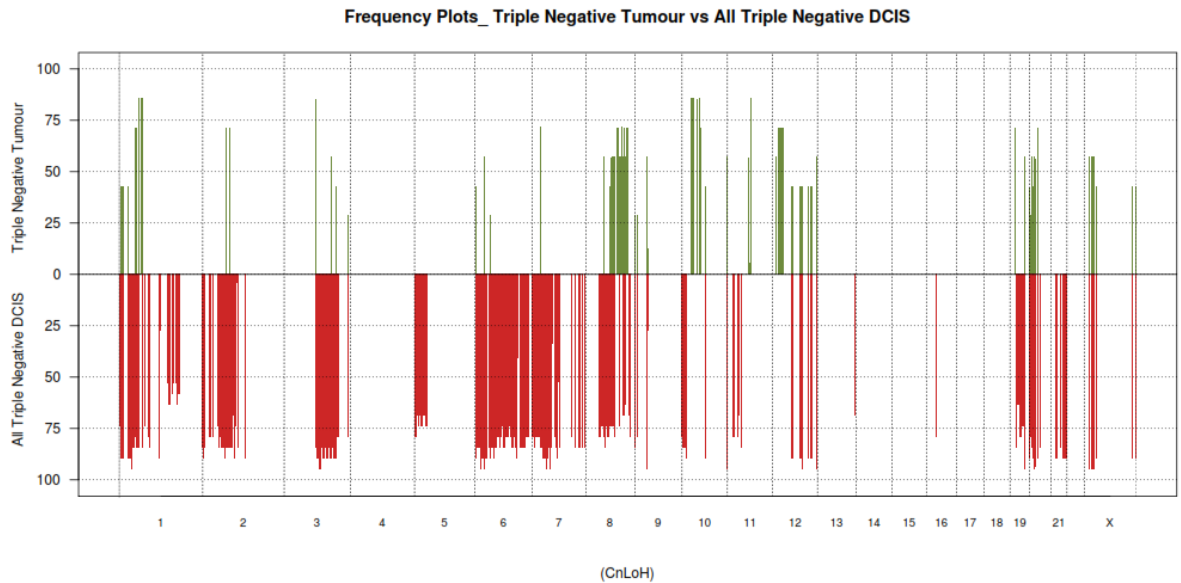
Frequency Plots_ Triple Negative Tumour vs All Triple Negative DCIS











5.8.10.1 Amplification of Triple Negative DCIS and Triple negative Invasive Breast Disease

No amplifications were present in triple negative DCIS (n=0/19) or triple negative invasive breast disease (n=0/7).

5.8.10.2 Duplication of Triple Negative DCIS Compared to Duplication in Triple Negative Invasive Breast Disease

Duplicated genes found in TN DCIS (1/19) were also present in TN invasive breast disease (n=3/7).

1. p-values < 0.05;
2. The gene list is created from 5 genomic regions found on chromosomes 2, 6, 12;
3. These regions encompass 57 genes altered in TN DCIS;

Genes:

7SK	AC012494.1	AC079600.1	AL121959.1	ARG1	C6orf192	CTAGE9	CTGF	EEF1A1P36	ENPP1	ENPP3
EYA4	HCRT2	HMGB1L13	LRFN2	MED23	MOXD1	OR2A4	PRICKLE1	RP11-121P10.1	RP11-123H21.1	
RP11-203B4.1		RP11-25115.1			RP11-295F4.4		RP11-314E23.1	RP1-131F15.2	RP11-69I8.2	
RP11-69I8.3	RP1-283K11.3	RP1-55C23.4			RP1-55C23.5		RP1-55C23.7	RP1-78N10.2	RP3-	

323K23.3	RP3-416F10.1	RP3-462C17.1	RP3-523C21.1	RP3-523C21.2	RP5-988G15.1
RPS12	SNORA33	SNORD100	SNORD101	STX7	TAAR1
TAAR7P	TAAR8	TAAR9	U4	VNN1	VNN2
				VNN3	

PANTHER analysis: 22 mapped ids are found, 35 mapped ids are not found.

Duplication of genes found in TN invasive breast disease (n=3/7) which are not observed in TN DCIS (n=0/19):

1. p-values < 0.05;
2. The gene list is created from 8 genomic regions found on chromosomes 6, 8, 12;
3. These regions encompass 126 genes altered in TN invasive breast disease;

Genes:

5S_rRNA	7SK	ABCC10	AL080315.3	AL096711.1	AL136304.1	AL136304.2	AL136304.3	AL451073.1	AL583834.1	BMP5
BYSL		C6orf108	C6orf132	C6orf15_2	C6orf154	C6orf174	C6orf226	C6orf58	C8orf85	CCND3
CENPW	CLIC5	CNPY3								
COL21A1	CRIP3	CUL7	CUL9	DLK2	ECHDC1	EEF1DP5	ENPP4	ENPP5	FRS3	GFRAL
GNMT	GUCA1A	GUCA1B	HCRT2	HEY2						
HINT3	HMGCLL1	hsa-mir-588	KIAA0240	KLC4	KLHDC3	LAMA2	MEA1	MED20	MRPL2	MRPS10
NCOA7	PEX6	PGC								
POLR1C	PPP2R5D	PRB2	PRICKLE4	PTCRA	PTK7	PTPRK	RNF146	RP11-103C16.2	RP11-151M7.1	
RP11-162L10.1		RP11-213N20.1		RP11-228O6.2		RP11-298J23.5	RP11-298J23.6	RP11-325O24.1	RP11-325O24.2	RP11-357P24.2
RP11-139D8.6	RP11-1480N24.3	RP11-480N24.4	RP11-480N24.6	RP11-527F13.1	RP11-527F13.2	RP11-533O20.2	RP11-570K4.1	RP11-624M8.1	RP11-753G20.1	RP11-775I22.2
RP1-177A13.1	RP1-179E13.1	RP1-293L8.2								
RP1-86D1.2	RP1-86D1.3	RP1-86D1.4	RP1-86D1.5	RP1-8B1.4	RP3-330M21.5	RP3-337H4.6	RP3-351K20.3	RP3-351K20.4	RP3-403A15.1	RP3-447E21.3
RP3-475N16.1	RP5-973N23.4									
RPL24P4	RPL7L1	RSPO3	SCARNA15	SLC22A7	SLC30A8	SNORA8	snoU13	SRF	TAF8	TFEB
THEMIS										
TJAP1	TOMM6	TRERF1	TRMT11	TTBK1	U6	U7	USP49	XPO5	Y_RNA	YIPF3
ZNF318										

PANTHER analysis: 62 mapped ids are found, 64 mapped ids are not found.

5.8.10.3 Genomic Gains of Triple Negative DCIS Compared to Genomic Gains in Non Triple Negative DCIS

Gene gains for TN DCIS (n= 4/19) and TN invasive breast disease (n=5/7) show considerable overlap, i.e. genes are present in both groups.

1. p-values < 0.05;
2. A gene list is mapped from 25 genomic regions found on chromosomes 1, 3, 6, 9, 10, 20, 21, X;

3. These regions encompass 232 genes altered in TN DCIS and TN invasive breast disease;

Genes:

5S_rRNA 7SK AC055758.1 AC072028.2 AC092898.1 AC107021.1 AC108729.1 AC112504.2 AC128648.1 AC133435.1
AC133435.2 AKAP12 AL096840.2 AL133260.1 AL158155.2 AL356137.1 AL357514.2 AL357561.2 AL365214.2 AL450307.1
AL583849.1 ATP1B3 ATR BNIP3 C1orf135 C3orf16 C3orf58 C3orf70 CCDC39 CCIN CHCHD6 CLIC6 CLTA COMMD2
CTD-2501O3.1 CTD-2501O3.2 CTD-2501O3.3 DCAF10 DHX36 EEF1A1P25 EHHADH EIF2S2P2 ESR1 EXOSC3 EXTL1 FAM54B
FBXO10 FLYWCH1P1 FRMPD1 FXR1 GK5 GLIPR2 GM2AP1 GNE GPR149 GRHPR GRK7 HRCT1 IPCEF1 IYD JAKMIP3
MACROD2 MAN1C1 MAP3K13 MELK MFN2 MIIP MME MTHFD1L NCRNA00160 OPRM1 OR2S2 PABPC1P10
PAFAH2 PAQR7 PAX5 PCOLCE2 PDIK1L PFN2 PIK3CB PLCH1 PLEKHG1 PLOD2 PLS1 PLSCR1 PLSCR2 PLSCR4 PLSCR5 PLXNA1
POLR1E PPP1R14C PPP2R2D RAB1C RECK RFPL4B RG9MTD3 RNF13 RNF38 RNF7 RP1-105O18.1 RP11-
110C15.4 RP11-113A10.1 RP11-113A10.2 RP11-113A10.3 RP11-113A10.5 RP11-117F22.1 RP11-117L15.1
RP11-136K14.1 RP11-136K14.2 RP11-136K14.3 RP11-144C9.1 RP11-167H9.2 RP11-167H9.3
RP11-167H9.4 RP11-167H9.5 RP11-167H9.6 RP11-190C21.3 RP11-190C21.4 RP11-217E22.1
RP11-217E22.2 RP11-217E22.5 RP11-21M4.2 RP11-220I1.1 RP11-220I1.2
RP11-220I1.4 RP11-232M24.1 RP1123D24.1 RP11-23D24.2 RP11-255N4.2 RP11-255N4.3
RP11-259P15.1 RP11-259P15.3 RP1-125I3.2 RP1-125I3.3 RP1-125I3.4 RP1-125I3.7 RP11-270G15.2 RP11271K21.2
RP11-271K21.4 RP11-271K21.7 RP11-274H2.2 RP11-274H2.3 RP11-291C6.1 RP11-297B17.2 RP11-29A1.3
RP11-317B3.2 RP11-327L3.1 RP11-327L3.3 RP11-327L3.4 RP11-327L3.5
RP11-340E6.1 RP11-340E6.2 RP11-343B5.1 RP11-383G6.3 RP11-383G6.4 RP11-397D12.4
RP11-397D12.6 RP11-397D12.7 RP11-405L18.1 RP11-405L18.2 RP11-405L18.4 RP11427D1.1 RP11-439C8.1
RP11-439C8.2 RP11-451G4.1 RP11-451G4.2 RP11-451G4.3 RP11-465M18.1 RP11-
471B18.1 RP11-483E7.1 RP11-506B6.1 RP11-506B6.2 RP11-506B6.3 RP11-548O1.1 RP1-159M24.1
RP11-605F14.3 RP11-613F7.1 RP11-613M10.6 RP11-622L21.1 RP11639F1.1 RP11-639F1.2
RP11-639F1.3 RP11-63E16.1 RP11-649A16.1 RP11-651P23.2 RP11651P23.3
RP11-651P23.4 RP11-651P23.5 RP11-651P23.6 RP11-656A15.1 RP11-65C6.1
RP11-758I14.2 RP11-758I14.3 RP11-785G3.1 RP11-84P7.1 RP11-84P7.2
RP11-84P7.3 RP11-88H10.1 RP1188H10.2 RP11-88H10.3 RP1-197O17.3 RP1-292B18.1 RP1-
292B18.3 RP1-292B18.4 RP1-297M16.2 RP1-317E23.3 RP1-44A20.4 RP1-99E18.2 RP3-422F24.2 RP3-
460G2.2 SCARNA17 SCARNA18 SEPNI1 SLC30A2 SLC35F1 snoU13 STMN1 TFDP2 TOMM5 TRIM63 TRPC1 U1 U3 U4 U6 U7 U8
WWTR1 XRN1 Y_RNA ZBTB5 ZCCHC7

PANTHER analysis: 77 mapped ids are found, 155 mapped ids are not found.

5.8.10.4 Sc Gains of Triple Negative DCIS Compared to Sc Gains in Triple Negative Invasive Breast Disease

The comparison of values for Sc gains in TN invasive breast disease (n=3/7) and TN DCIS (n=0/19)

1. p-values <0.05;
2. A gene list is mapped from 32 genomic regions found on chromosomes 1, 2, 3, 6, 10, 16, 20. This represents Sc gains for TN Invasive breast disease;

3. These regions encompass 209 genes altered in TN invasive breast disease;

Genes:

5S_rRNA 7SK AADAC AADACL2 AC007392.3 AC007403.1 AC007403.2 AC009086.1 AC009086.3 AC009086.4
AC009474.1 AC009474.2 AC025279.1 AC055758.1 AC072028.2 AC078784.1 AC107021.1 AC107027.2 AC108729.1
AC121332.1 AC128648.1 AC133435.1 AC133435.2 AC133555.1 AC133555.2 AC133555.3 AC133555.4 AC133555.6 AC133555.7
ACPL2 AL357561.2 AL450307.1 ARHGAP31 ASTE1 ATP2C1 ATP5LP5 ATR B4GALT4 BNIP3 BOLA2
C1orf135 C2orf86 C3orf16 C3orf30 C3orf58 CCDC39 COMMD2 CPNE4 CTD-2501O3.1 CTD-2501O3.2
CTD-2501O3.3 DHX36 EEF1A1P25 EXTL1 FAM54B GIYD1 GK5 GM2AP1 GPR149 GSTO3P IGSF11 IYD JAKMIP3
LSAMP MACROD2 MAN1C1 MME MRPL3 NEK11 NUDT16 NUDT16P OPRM1 P2RY1 PABPC1P10
PAFAH2 PAQR7 PARK2 PCOLCE2 PDIK1L PFN2 PIK3CB PLCH1 PLOD2 PLS1 PLSCR1
PLSCR2 PLSCR4 PLSCR5 PPP1R14C PPP2R2D QPRT RASA2 RNF13 RP11-117F22.1
RP11-117L15.1 RP11-167H9.2 RP11-167H9.3 RP11167H9.4 RP11-167H9.5 RP11-167H9.6
RP11-190C21.3 RP11-190C21.4 RP11-217E22.1 RP11217E22.2 RP11-217E22.5 RP11-21M4.2 RP11-
232M24.1 RP11-23D24.1 RP11-23D24.2 RP11255N4.2 RP11-255N4.3 RP1-125I3.2 RP1-125I3.3 RP1-125I3.4 RP1-
125I3.7 RP11265F19.1 RP11-270G15.2 RP11-274H2.2 RP11-274H2.3 RP11-291C6.1 RP11-305I9.1
RP11-316C10.1 RP11-340E6.1 RP11-343B5.1 RP11-383G6.3 RP11-383G6.4
RP11-38P22.2 RP11-39E3.3 RP11-39E3.4 RP11-39E3.5 RP11-39E3.6 RP11-402E20.1
RP11-420J11.2 RP11-427D1.1 RP11-438D8.2 RP11-438D8.3 RP11-438D8.4 RP11-439C8.1
RP11-439C8.2 RP11-451G4.1 RP11-451G4.2 RP11-451G4.3 RP11-454C18.2 RP11-483E7.1
RP11-484M3.4 RP11-484M3.5 RP11-496B10.1 RP11-496B10.3 RP11-496B10.5
RP11-517B11.2 RP11-517B11.4 RP11-517B11.6 RP11-548O1.1 RP11-622L21.1
RP11-639F1.1 RP11-639F1.2 RP11-639F1.3 RP11-63E16.1 RP11-63L4.1 RP11-649A16.1
RP11-64D22.1 RP11-64D22.2 RP11-64D22.5 RP11-651P23.2 RP11-651P23.3
RP11-651P23.4 RP11-651P23.5 RP11-651P23.6 RP11-656A15.1 RP11-758I14.2
RP11-758I14.3 RP11-785G3.1 RP11-788A4.1 RP11-788A4.2 RP11-789L4.1
RP11-88H10.1 RP11-88H10.2 RP11-88H10.3 RP11-933H2.4 RP1-317E23.3 RP3-422F24.2 RP4-635B5.1
RUNDC2B SCARNA17 SCARNA18 SEPN1 SLC30A2 SNORA58 snoU13 SPN STMN1 SUCNR1 SULT1A4 TFDP2 TM4SF4
TRIM63 TRPC1 TTC14 U1 U3 U4 U6 U7 U8 UPK1B WWTR1 XRN1 Y_RNA
ZBTB38

PANTHER analysis: 64 mapped ids are found, 145 mapped ids are not found.

5.8.10.5 Losses in Triple Negative DCIS Compared to Losses in Triple Negative Invasive Breast Disease

Losses in TN DCIS (n=1/19) show similarities with losses in TN invasive breast disease (n=3/7). This represents losses for TN DCIS and TN invasive breast disease.

1. p-values <0.05;
2. The gene list is created from 31 genomic regions found on chromosomes 1, 2, 3, 6, 10, 16, 20;
3. These regions encompass 203 genes altered in TN DCIS;

Genes:

5S_rRNA	7SK	AADAC	AADACL2	AC007392.3	AC007403.1	AC007403.2	AC009086.1	AC009086.3	AC009086.4	
	AC009474.1	AC009474.2	AC025279.1	AC055758.1	AC072028.2	AC078784.1	AC107021.1	AC107027.2	AC108729.1	AC121332.1
	AC133555.1	AC133555.2	AC133555.3	AC133555.4	AC133555.6	AC133555.7	ACPL2	AL357561.2	AL450307.1	ARHGAP31
ASTE1	ATP2C1	ATP5LP5	ATR	B4GALT4	BNIP3	BOLA2	C1orf135	C2orf86	C3orf16	C3orf30
	C3orf58	CCDC39	COMMD2	CPNE4	CTD-2501O3.1	CTD-2501O3.2	CTD-2501O3.3		DHX36	
	EEF1A1P25	EXTL1	FAM54B	GIYD1	GM2AP1	GPR149	GSTO3P	IGSF11	IYD	JAKMIP3
MACROD2	MAN1C1	MME	MRPL3	NEK11	NUDT16	NUDT16P	OPRM1	P2RY1	PABPC1P10	
	PAFAH2	PAQR7	PARK2	PCOLCE2	PDIK1L	PFN2	PIK3CB	PLCH1	PLOD2	PLS1
	PLSCR1	PLSCR2	PLSCR4	PLSCR5	PPP1R14C	PPP2R2D	QPR1	RASA2	RNF13	RP11-117F22.1
	RP11-117L15.1		RP11-167H9.2		RP11-167H9.3	RP11-167H9.4		RP11-167H9.5		RP11-167H9.6
	RP11-190C21.3		RP11-190C21.4		RP11-217E22.1		RP11-217E22.2		RP11-217E22.5	
	RP11-21M4.2		RP11-232M24.1		RP11-23D24.1		RP11-23D24.2		RP11-255N4.2	
	RP11-255N4.3		RP1-125I3.2	RP1-125I3.3	RP1-125I3.4	RP1-125I3.7	RP11-265F19.1		RP11-270G15.2	
	RP11-274H2.2		RP11-274H2.3		RP11-291C6.1		RP11-305I9.1		RP11-316C10.1	
	RP11-340E6.1	RP11-383G6.3		RP11-383G6.4		RP11-38P22.2		RP11-39E3.3		RP11-39E3.4
	RP11-39E3.5		RP11-39E3.6		RP11-402E20.1		RP11-420J11.2		RP11-427D1.1	
	RP11-438D8.2		RP11-438D8.3		RP11-438D8.4		RP11-439C8.1		RP11-439C8.2	
	RP11-451G4.1		RP11-451G4.2		RP11-451G4.3		RP11-454C18.2		RP11-483E7.1	
	RP11-484M3.4		RP11-484M3.5		RP11-496B10.1		RP11-496B10.3		RP11-496B10.5	
	RP11-517B11.2		RP11-517B11.4		RP11-517B11.6		RP11-548O1.1		RP11-622L21.1	
	RP11-639F1.1		RP11-639F1.2		RP11-639F1.3		RP11-63E16.1		RP11-63L4.1	RP11-649A16.1
	RP11-64D22.1		RP11-64D22.2		RP11-64D22.5	RP11-651P23.2		RP11-651P23.3		RP11-651P23.4
	RP11-651P23.5	RP11-651P23.6		RP11-656A15.1		RP11-758I14.2		RP11-758I14.3		RP11-785G3.1
	RP11-788A4.1		RP11-788A4.2		RP11-789L4.1		RP11-88H10.1	RP11-88H10.2		RP11-88H10.3
	RP11-933H2.4		RP1-317E23.3	RP3-422F24.2		RP4-635B5.1		RUNDC2B	SCARNA17	SCARNA18
SEPN1	SLC30A2	SNORA58		snoU13	SPN	STMN1	SUCNR1	SULT1A4	TM4SF4	TRIM63
TRPC1	TTC14	U1	U3	U4	U6	U7	U8	UPK1B		WWTR1
XRN1	Y_RNA	ZBTB38								

PANTHER analysis: 62 mapped ids are found, 141 mapped ids are not found.

The comparison of losses represents losses in TN DCIS and TN invasive breast disease

1. p-values < 0.05;
2. A gene list is mapped from 98 genomic regions found on chromosomes 1, 4, 10, 12, 15, 18, 19, X;
3. These regions encompass 317 genes altered in TN DCIS and TN invasive breast disease;

Genes:

5S_rRNA	AC005184.1	AC008245.1	AC009731.1	AC012379.1	AC012379.2	AC013452.1	AC013452.2	AC013558.1	AC018413.1	
	AC018797.3	AC022046.2	AC025263.1	AC034110.1	AC034110.2	AC036176.2	AC064801.1	AC068473.1		AC083805.1
	AC083805.2	AC090360.1	AC090393.1	AC090669.1	AC091551.1	AC091643.1	AC091646.1	AC093827.1	AC100783.1	
	AC100843.1	AC104059.1	AC114688.1	AC134978.1	AC139100.1	AC139100.2	ACE2	ADNP2		AF196969.5
AF196972.1	AF196972.10		AF196972.2	AF196972.3	AF196972.4	AF196972.9	AF213884.1	AF224669.3	AF241726.2	

AF241726.4	AF241726.6	AFF1	AL031311.1	AL136137.1	AL353691.1	AL356741.1	AL591378.1	AL591394.1	
AL592043.1	AL627402.1	ALPK2	ANXA3	APOF	ARAF	ARD1B	ARHGAP24	ARHGEF16	ATP9B
AVIL	BCL2	BEST3	BMX	C18orf22	C4orf36	CA5B	CA5BP	CASK	CBLN2
CDH19	CDH7	CDK4	CEP152	CHST7	CHTF8P	COPS2	CTD-2522E6.4	CTDP1	
CTDSP2	CXorf24	CXorf36	CXorf58	CYP27B1	CYP2A6	DNAJC12	DUSP21	EBP	
EIF2S3	ELAC1	FAM47A	FGF5	FIGF	FRAS1	FTLP16	FTSJ1	GALK2	
GALR1	GPR34	GPR82	GRPR	GS1-541M1.1		GS1-541M1.2		GS1-594A7.3	
GS1-594A7.5		HERC4	HMG1L1	hsa-mir-221	hsa-mir-222	hsa-mir-26a-2		hsa-mir-551a	
HSBP1L1	IL1RAPL1	INE1	KAL1	KCNG2	KDM6A	KDSR	KIAA1468	KLHL15	KLHL8
KRT8P14	L1CAM	MALT1	MANBA	MAOA	MAOB	MAPK10	March9	MBP	ME2
MEGF6	MEX3C	NDP	NDUFB11	NFATC1	NFKB1	NTS	NUAK1	NUS1P1	OTC
PARD6G	PCTK1	PIGN	PIR	PORCN	PPP1R2P9	PQLC1	PRDM8	PTCHD2	
PTPN13	RAB3IP	RASGEF1B	RASSF9	RBM10	RBM19	RBM3	RNF152	RP11-10L12.1	
RP11-10L12.2		RP11-10L12.4		RP11-10L12.5		RP11-10L12.6		RP11-1148L6.5	
RP11-1148L6.6		RP11-126O1.1		RP11-149B9.2		RP11-162A12.1	RP11-162K6.1	RP11-204C16.4	
RP11-245M24.1		RP11-281B1.2		RP11-305F18.1		RP11-342D14.1		RP11-344N17.11	
RP11-344N17.12		RP11-344N17.13		RP11-344N17.2		RP11-344N17.3		RP11-344N17.6	
RP11-344N17.7		RP11-344N17.8		RP11-344N17.9		RP11-377B2.2		RP11-377G16.2	
RP11-38O23.3		RP11-38O23.4		RP11-38O23.5		RP11-38O23.7		RP11-397E7.1	
RP11-397E7.2		RP11-397E7.3		RP11-397E7.4		RP11-438E5.1		RP11-474D14.2	
RP11-476C8.2		RP11-479I1.4		RP11-524P6.1		RP11-540L11.2		RP11-545D19.1	
RP11-154K9.2		RP11-552E4.2		RP11-552E4.3		RP11-552E4.4		RP11-552E4.5	
RP11-571E6.1		RP11-571E6.3		RP11-571E6.4		RP11-57G10.6		RP11-617O8.1	
RP11-640A1.1		RP11-644A7.1		RP11-652N17.1		RP11-689K5.3		RP11-692P14.1	
RP11-702C7.1		RP11-702C7.2		RP11-75A9.2		RP11-75A9.3		RP11-778J15.1	
RP11-792D21.1	RP11-792D21.2		RP11-9H16.1		RP1-22N22.1		RP1-290C9.2		
RP13-202B6.2		RP13-479F17.2		RP13-928P6.3		RP1-50A13.1		RP1-50A13.2	
RP1-54B20.7		RP1-69M21.2		RP1-71L16.1	RP1-71L16.2	RP1-71L16.6	RP3-326I13.1	RP3-	
393P12.2	RP3-393P12.3	RP4-733D15.1		RP5-1174J21.1		RP5-1174J21.2	RP5-972B16.2	RP6-105D16.1	
RP6-218J18.3		RP6-99M1.1	RPGR	RPL12P8	RTTN	SALL3	SERPINB5	SHC4	SIRT1
SLC10A6	SLC38A5	SLC9A7	SMAD4	SNORA11		SNORA25	SNORA31	SNORA68	SNORD56
snoU13	SOCS6	SPACA5	SPACA5B	SSX1	SSX3	SSX4	SSX4B	SSX5	SSX6
SSX9	STAT2	SYN1		TBC1D25	TIMP1	TMEM27	TNFRSF11A	TSHZ1	TSPAN7
TXNL1	TXNL4A	U1	U2	U52112.12	U6	U7	U8	UBA1	
UBE2D3	USP11	VCX3B	VENTXP1	WAS	WDR13	WDR7	Y_RNA	Z98304.1	ZADH2
ZFX	ZNF157	ZNF182	ZNF236	ZNF407	ZNF41	ZNF516		ZNF532	ZNF630
ZNF673	ZNF674	ZNF81							

PANTHER analysis: 138 mapped ids are found, 179 mapped ids are not found.

5.8.10.6 *Total Loss of Triple Negative DCIS Compared to Total Loss in Triple Negative Invasive Breast Disease*

There were total losses present in TN DCIS (n=13/19) that are not observed from TN invasive breast disease (n=0/7).

1. p-values < 0.05;

2. A gene list is mapped from 150 genomic regions found on chromosomes 3, 4, 5, 6, 8, 9, 13, 14, 15, 17, 18, 22 & X;
3. These regions encompass 419 genes altered in TN DCIS;

Genes:

5S_rRNA	7SK	AARSD1	AB019438.55	AB019438.63	AB019438.66	AB019439.3	ABCD4					
	AC002117.1	AC003043.1	AC003043.2	AC003102.3	AC003963.2	AC004222.1	AC004596.1	AC004968.1	AC004968.2	AC005399.1		
	AC006059.1	AC006349.1	AC006445.1	AC006445.6	AC006445.7	AC006445.8	AC007126.1	AC007722.1	AC008105.1	AC011193.1		
	AC012409.1	AC015936.1	AC015936.3	AC016251.1	AC019131.2	AC020704.1	AC021205.1	AC025287.1	AC025287.2	AC025287.3		
AC025287.4	AC026202.1	AC026202.5	AC027139.2	AC027139.3	AC027139.4	AC027708.1	AC027807.1	AC027807.2	AC027807.3	AC036222.1		
	AC044797.1	AC044797.2	AC044797.3	AC055813.1	AC055873.1	AC068400.1	AC079915.1	AC084809.2	AC084809.3	AC084882.1		
AC087650.1	AC087742.1	AC092047.1	AC092048.1	AC092421.1	AC098830.1	AC100793.2	AC102948.1	AC102948.2	AC102948.3	AC103702.1		
AC103965.1	AC104260.1	AC105750.1	AC107982.4	AC109357.1	AC112693.1	AC113171.1	AC114477.1	AC121253.1	AC124789.1			
	AC134669.1	AC138150.3	AC138150.4	AC183087.1	AC183087.2	ACBD4	ACTR10	ADAM11	ADAM3A	ADAM5P	ADAMTSL3	
ADH4	ADH5	AGBL1	AKAP6	AL008582.1	AL035659.2	AL035659.4	AL035659.6	AL079307.3	AL079307.4			
AL079307.5	AL132989.2	AL135978.1	AL136359.1	AL139021.2	AL139089.1	AL162377.1	AL162377.2	AOC2	AOC3	AP000351.4		
ARAP2	ARHGAP23	ARID4A	ARL4D	ARRDC4	ASB16	ASPN	ATXN7L3	AY269186.1	AY269186.2	AY269186.3	AZI2	
BBS12	BOD1L	BRCA1	BTF3	C14orf149	C14orf166	C14orf38	C14orf54	C17orf102	C17orf104			
	C17orf105	C17orf46	C17orf53	C17orf65	C1QL1	C4orf37	CA10	CARTPT	CCDC103	CCDC43	CD300LG	CDC6
CECR2	CECR9	CENPP	CETN4P	CHADL	CMC1	CSMD1	CTD-2015H6.1	CTD-2015H6.2	CTD-2015H6.3	CTNS	CWC25	
	DBF4B	DCAKD	DDT	DDX52	DHRS11	DHX8	DTHD1	DUSP3	ECM2	EDEM1	EFTUD2	EIF4BP4
	EIF4E	EP300	ETV4	EVI2A	EVI2B	FAM149B2	FAM151B	FAM169B	FAM171A2	FAM177A2		
	FAM187A	FAM18B	FAM19A1	FBXL2	FBXO33	FBXW10	FGF2	FLT3	FMNL1	FOXD1	FRMD6	
FZD2	G6PC	G6PC3	GFAP	GJC1	GNG2	GOLGA8B	GPATCH8	GPC5	GPHN	GPR179	GRN	
	GSTT2	HDAC5	HEXIM1	HEXIM2	HIGD1B	HMGB1L14	HOXB7	HOXB9	hsa-mir-1233	hsa-mir-1281	hsa-mir-1469	hsa-mir-196a-1
mir-196a-1	hsa-mir-2117	hsa-mir-759	HSP90AB2P	IFI35	IGHV1-45	IGHV1-46	IGHV3-33-2	IGHV3-35	IGHV3-36			
	IGHV3-37	IGHV3-38	IGHV3-41	IGHV3-42	IGHV3-43	IGHV3-47	IGHV4-34	IGHV4-39	IGHV7-34-1	IGHV7-40	IGHVII-33-1	IGHVII-40-1
IGHVII-40-1	IGHVII-43-1	IGHVII-44-2	IGHVII-46-1	IGHVIII-38-1	IGHVIII-44	IGHVIV-44-1	IKZF3	IPO11	ITGA2B	ITGAE		
JKAMP	KCNH5	KIAA0586	KIF18B	KIF2A	KLHL5	L3MBTL2	LECT1	LIN52	LPAR4	LRRC70		
LSM12	LYZL4	MAP3K14	MCTP2	MEOX1	METAP1	MITF	MPP2	MPP3	MPP5	MRM1		
MRPL45	MYO1D	NAGS	NF1	NID2	NKTR	NKX3-2	NMT1	NR2F2	NUDT6	OCN2	OLFM4	OMD
OMG	P2RX5	PARK2	PCDH8	PDCD6IP	PDS5B	PIP5K2B	PLCD3	PPY	PRPSAP2	PSMA3	PSMB3	PSME3
PYY	RAB28	RAD51L1	RANGAP1	RGS6	RHCGRNF180	RP11-1072C15.1	RP11-1280N14.3	RP11-1280N14.4				
	RP11-1299A16.1	RP11-1299A16.2	RP11-12M9.4	RP11-159J2.1	RP11-159J2.2	RP11-170N16.1						
	RP11-170N16.2	RP11-187O7.1	RP11-209K10.1	RP11-219I21.1	RP11-223D18.1							
	RP11-241P3.2	RP11-248N6.1	RP11-248N6.2	RP11-24H2.1	RP11-24H2.2	RP11-301J16.5						
	RP11-30C8.1	RP11-30C8.2	RP11-327P2.5	RP11-341G5.1	RP11-357H14.4	RP11-360F5.1						
	RP11-360F5.2	RP11-360F5.3	RP11-369I16.1	RP11-380B4.2	RP11-428B4.2	RP11-431M7.2						
	RP11-431M7.3	RP11-431O22.2	RP11-442J17.4	RP11-474L7.4	RP11-475D8.1	RP11-48B14.1						
	RP11-520F24.1	RP11-520F24.2	RP11-520F24.3	RP11-522B15.1	RP11-522B15.2	RP11-571L19.2						
	RP11-571L19.4	RP11-696N14.1	RP11-799A12.1	RP11-799A12.2	RP11-79P5.2	RP11-79P5.3	RP11-79P5.5					
	RP11-79P5.6	RP11-79P5.7	RP11-79P5.8	RP11-81N13.1	RP11-85I21.1	RP11-880O3.1						
	RP11-95I19.1	RP11-95I19.2	RP1-85F18.4	RP1-85F18.5	RP1-85F18.6	RP4-613B23.1	RP4-756G23.5	RPL27	RTN1			
RUNDC1	RUNDC3A	SEC22C	SHISA6	SHPK	SIPA1L1	SLC25A21	SLC25A39	SLC4A1	SMN1	SMOC1		
SNORA18	SNORA40	SNORD56	SNORD74	snoU13	snoZ178	SOCS7	SOST	SPATA5	SPATA8	SPRED1	SRCIN1	
SS18L2	SUCLG2	SYNRG	SYT16	TAX1BP3	TBC1D3	TEF	TIMM9	TMEM101	TMEM132E	TMEM93		

TMEM98	TMUB2	TOMM20L	TRPV1	TRPV3	TSPAN5	U3	U6	U7	UBP1	UBTF	VAT1	VSX2
WIPF2	WSCD1	Y_RNA	ZBTB47	ZC3H7B	ZCCHC5	ZCWPW2	ZFYVE16					

PANTHER analysis: 194 mapped ids are found, 225 mapped ids are not found.

There were total losses present in TN invasive breast (n=7/7) disease that are not observed in TN DCIS (n=0/19).

1. p-values < 0.05;
2. A gene list is mapped from 98 genomic regions found on chromosomes 3, 4, 5, 9, 10, 12, 13, 14, 15, 17;
3. These regions encompass 479 genes altered in TN invasive breast disease;

Genes:

5S_rRNA	7SK	AATF	AB019437.56	ABHD12B	ABLM2	ABR	AC003664.1	AC005277.1	AC005304.1
	AC005304.2	AC005304.3	AC005324.1	AC005324.2	AC005324.4	AC005324.6	AC005324.7	AC005358.1	AC005358.3
AC005517.2	AC005517.3	AC005703.2	AC005703.3	AC005772.2	AC005773.1	AC005838.2	AC005863.1	AC005863.2	AC007686.1
	AC013248.1	AC013248.2	AC015842.1	AC015922.1	AC015922.2	AC015922.7	AC024610.1	AC027045.1	AC073940.1
	AC087742.2	AC090283.1	AC090283.3	AC090516.1	AC090516.2	AC090519.1	AC090519.10		AC090519.11
	AC090519.2	AC090519.3	AC090519.4	AC090519.5	AC090519.6	AC090519.7	AC090519.8	AC090519.9	AC091487.1
	AC093824.1	AC098487.1	AC103560.1	AC104650.2	AC108477.2	AC123769.1	ACCN1	ADCK1	AE000658.6
AE000659.10	AE000659.12		AE000659.13		AE000659.14		AE000659.15		AE000659.16
AE000659.17	AE000659.18		AE000659.19		AE000659.23		AE000659.6	AE000659.8	AE000659.9
	AF111168.1	AF111168.3	AGPAT9	AHSA1	AL049775.1	AL121768.1	AL133223.1	AL133413.1	AL136160.1
	AL139041.1	AL157702.1	AL158058.1	AL162727.1	AL353141.1	AL353782.1	AL353805.1	AL354751.2	AL354751.3
	AL356214.2	AL357172.1	AL358334.1	AL358334.2	AL359180.1	AL359392.1	AL360004.1	AL365364.1	AL732479.1
	ALKBH1	AMBP	ANKFY1	ANKRD40	ANKS6	ARD1B	ASTN2	ATAD1	ATP5C1P1
	C10orf4	C13orf23	C13orf36	C14orf148	C14orf156	C14orf174	C14orf178	C14orf184	C14orf23
	C9orf109	C9orf110	C9orf125	C9orf147	C9orf170	C9orf30	C9orf44	C9orf80	CACNA1G
	CDC20P	CDKN2AIP	CDRT1	CDRT4	CDRT7	CEP55	CNOT6L	COL27A1	CTC-448D22.1
	CTC-507E12.1		CTD-2158P22.1		CTD-2158P22.2		CTD-2158P22.4		CTD-2175M1.1
	CTDSPL2	CYB5D2	CYP26A1	CYP26C1	DAB2IP	DBC1	DDIT4L	DHRS7C	DKK2
	DLEU2	DTWD2	DYNLL1P1	ELAC2	EML5	EXOC6	FAM154A	FAM170A	FAM175A
	FAM190	FAM190A	FAM33B	FKBP15	FLRT2	FOXG1	FRMD5	FYTDT1P1	GABBR2
	GALNTL6	GAPDHP34	GAS1	GAS7	GK2	GLP2R	GMPSP	GNG5P5	GPM6A
	GRIN3A	GSTZ1	HELT	HS3ST3A1	hsa-mir-1260	hsa-mir-15a	hsa-mir-16-1	hsa-mir-346	hsa-mir-455
	hsa-mir-95	HSD17B4	HSDL2	HTRA3	IGHV1-67	IGHV1-68	IGHV1-69	IGHV2-70	IGHV3-62
	IGHV3-65	IGHV3-66	IGHV3-71	IGHV3-72	IGHV3-73	IGHV4-61	IGHVII-60-1	IGHVII-62-1	IGHVII-65-1
	IGHVIII-67-2	IGHVIII-67-3	IGHVIII-67-4	IL6STP1	ING2	ISM2	KIAA1737	KIAA1958	KIF12
	KRT25	KTN1	LGI1	LHFP	LHX6	LMO7	MEIS3P1	MITF	MMRN1
	MRPL50	MRRF	MURC	MYOCD	MYOF	NDUFA8	NGB	NHLRC3	NRXN3
	OR1B1	OR1H1P	OR1J2	OR1J4	OR1L1	OR1L3	OR1L4	OR1L6	OR1L8
	OR1N2	OR1Q1	OR7E94P	PAPPA	PAPSS1	PAPSS2	PDE6C	PDZRN3	PMP22
	RASGEF1B	RBM18	RBP4	RCVRN	RFXAP	RGS3	RNF20	ROR2	RP11-100G15.10
	RP11-100G15.3		RP11-100G15.4		RP11-100G15.5		RP11-100G15.6		RP11-100G15.7
	RP11-115D19.1		RP11-115G5.1		RP11-118F2.2		RP11-118F2.3		RP11-122F14.4
1258F18.1		RP11-127L21.1		RP11-127L21.2		RP11-15B17.1		RP11-162D16.4	
									RP11-168K11.2

RP11-168K11.3	RP11-16L6.3	RP11-181F12.1	RP11-18B16.1	RP11-18B16.2	RP11-197L7.2
RP11-20O12.1	RP11-226P1.1	RP11-234K19.1	RP11-234K19.2	RP11-24E1.1	
RP11-264C15.2	RP11-276E15.4	RP11-276H19.1	RP11-276H19.2	RP11-276H19.4	RP11-280G19.1
RP11-280G19.2	RP11-296L20.1	RP11-299L17.1	RP11-32M23.4	RP11-332E3.2	
RP11-348J12.2	RP11-349E4.1	RP11-34F20.4	RP11-34F20.5	RP11-352E6.1	
RP11-352E6.2	RP11-35N6.1	RP11-35N6.3	RP11-35N6.4	RP11-35N6.6	RP11-367N14.1
RP11-367N14.2	RP11-385D13.1	RP11-386B13.1	RP11-386B13.3	RP11-386B13.4	RP11-388N2.1
RP11-394D2.1	RP11-395D3.1	RP11-397J18.1	RP11-408O19.3	RP11-421P11.5	
RP11-438E5.1	RP11-44F21.4	RP11-452B18.1	RP11-452B18.2	RP11-45A16.3	RP11-45A16.4
RP11-45I20.1	RP11-462C21.1	RP11-477L16.2	RP11-478K15.1	RP11-478K15.2	
RP11-478K15.3	RP11-478K15.4	RP11-478K15.5	RP11-480P3.1	RP11-490D19.6	
RP11-490D19.8	RP11-499E18.1	RP11-51N22.1	RP11-51N22.2	RP11-521H3.1	
RP11-524G24.2	RP11-527F15.1	RP11-534I8.1	RP11-542G1.1	RP11-542G1.2	
RP11-542G1.3	RP11-545P6.1	RP11-545P6.2	RP11-54K22.1	RP11-565F19.1	
RP11-567N4.2	RP11-567N4.3	RP11-576C12.1	RP11-588P8.1	RP11-58C3.2	
RP11-598G2.1	RP11-610O8.1	RP11-616K22.1	RP11-61P7.3	RP11-624A4.1	RP11-624L12.1
RP11-626E13.1	RP11-64P14.7	RP11-67K19.3	RP11-67M1.1	RP11-689K5.3	RP11-692E14.1
RP11-701P16.3	RP11-701P16.4	RP11-713M6.1	RP11-713M6.2	RP11-722P15.1	RP11-765K14.1
RP11-806K15.1	RP11-84H6.1	RP11-88G17.2	RP11-92C4.3	RP11-94C24.10	RP11-94C24.11
RPL10P3	RPS3AP5	RWDD4A	SEPT7P7	SH3TC1	SLC31A2
					SLC39A8
					SLC46A2
					SLC4A4
					SMAD9
					SMARCE1
SNCA	SNORA26	SNORA31	SNORA32	SNORA46	SNORA70
					SNORA74
					SNORD112
					SNORD65
					snoU13
snoU83B	SNW1	SNX30	SORBS2	SPRY2	SPTLC1
					SPTLC2
					STOML3
					STON2
					STOX2
					STX8
TBC1D26	TEKT3	TIMM22	TLK1P1	TMED8	TMEFF1
					TMEM20
					TMEM63C
					TRIM16
					TRIM32
TTC8	TTLL11	U1	U3	U4	U6
					U7
					UBE2G1
					USP43
					VEGFC
					VIPAR
WDR16	WDR20	Y_RNA	ZDHHC22	ZFP37	ZNF189
					ZNF286
					ZNF29P
					ZNF618
					ZNF883
ZYG11AP1	ZZEF1				

PANTHER analysis: 160 mapped ids are found, 319 mapped ids are not found.

5.8.10.7 CdLOH of Triple Negative DCIS Compared to CdLOH in Triple Negative Invasive Breast Disease.

There is considerable overlap between CdLOH in TN DCIS (n=1/19) and CdLOH in TN invasive breast disease (n=3/7).

1. p-values < 0.05;
2. A gene list is mapped from 55 genomic regions found on chromosomes 1, 10, 11, 12, 18, 19, X;
3. These regions encompass 133 genes altered in TN DCIS and TN invasive breast disease;

Genes:

Mar09	AC005184.1	AC009731.1	AC011260.2	AC022046.2	AC025263.1	AC036176.2	AC064801.1	AC083805.1	AC083805.2
	AC090666.1	AC090669.1	AC100783.1	AC100843.1	AC139100.1	AC139100.2	AL031311.1	AL136137.1	AL353691.1

AL356741.1	AL627402.1	ALPK2	ARHGEF16	AVIL	BCL2	BEST3	C10orf100	CASK	CBLN2	CD226
	CDK4	CHEK1	CHTF8P	CTDSP2	CXorf36	CXorf58	CYP27B1	CYP2A6	DNAJC12	DOK6
	DUSP21	EIF2S3	ELAC1	GPR34	GPR82	GRPR	GS1-541M1.1		GS1-541M1.2	
	HERC4	HMGA1L1	hsa-mir-26a-2	hsa-mir-551a	IL1RAPL1	KAL1	KDM6A	KDSR	KIAA1468	
KLHL15	L1CAM	MALT1	MAOA	MAOB	MBD2	ME2	MEGF6	NDP		NEDD4L
	NRG3	NTS	NUAK1	OS9	PARD6G	PHLPP1	PIGN	PPP1R2P9	RAB3IP	RASSF9
	RNF152	RP11-126O1.1	RP11-149B9.2	RP11-202D18.2		RP11-204C16.4		RP11-20E23.1		
	RP11-342D14.1	RP11-377B2.2		RP11-472N6.1		RP11-474D14.2		RP11-479I1.4		
	RP11-514F8.2	RP11-524P6.1	RP11-540L11.2	RP1-154K9.2		RP11-57G10.6		RP11-617O8.1	RP11-640A1.1	
	RP11-652N17.1		RP11-702C7.1	RP11-702C7.2		RP11-9H16.1		RP1-22N22.1		RP1-
290C9.2	RP1-50A13.1		RP1-50A13.2	RP3-326I13.1		RP5-1174J21.1		RP5-1174J21.2		RP6-
105D16.1	RP6-218J18.3		RPL12P8	RTTN	SCN8A	SERPINB5	SH2D4B	SIRT1	SNORA68	snoU13
TNFRSF11A	TXNL1	U1	U2	U52112.12	U6	U8	VCX3B	VENTXP1	VPS4B	WDR7
	ZCCHC2	ZFX	ZNF407	ZNF532						Y_RNA

PANTHER analysis: 65 mapped ids are found, 68 mapped ids are not found.

There are some similarities between CdLOH in TN invasive breast disease (n=4/7) and CdLOH in TN DCIS (n=1/19).

- 1 p-values < 0.05;
- 2 A gene list is mapped from 71 genomic regions found on chromosomes 1, 10, 11, 12, 18, 19, X;
- 3 These regions encompass 167 genes altered in TN DCIS and TN invasive breast disease;

Genes:

Mar09 5S_rRNA	AC005184.1	AC008245.1	AC009731.1	AC011260.2	AC022046.2	AC025263.1	AC036176.2	AC064801.1	AC083805.1
	AC083805.2	AC090360.1	AC090393.1	AC090666.1	AC090669.1	AC091551.1	AC091643.1	AC100783.1	AC100843.1
	AC114688.1								
	AC139100.1	AC139100.2	ADNP2	AF196972.3	AF196972.4	AL031311.1	AL136137.1	AL353691.1	AL356741.1
	AL627402.1								
ALPK2	ARHGEF16	ATP9B	AVIL	BCL2	BEST3	C10orf100	C18orf22	CASK	CBLN2
CD226	CDH19	CDH7	CDK4	CHEK1	CHTF8P	CTDSP2	CXorf36	CXorf58	CYP27B1
CYP2A6	DNAJC12	DOK6	DUSP21	EIF2S3	ELAC1	FTSJ1	GPR34	GPR82	GRPR
	GS1-541M1.2		HERC4	HMGA1L1	hsa-mir-26a-2	hsa-mir-551a	HSBP1L1	IL1RAPL1	KAL1
	KIAA1468	KLHL15	L1CAM	MALT1	MAOA	MAOB	MBD2	MBP	ME2
	NEDD4L	NFATC1	NRG3	NTS	NUAK1	OS9	PARD6G	PHLPP1	PIGN
	RAB3IP	RASSF9	RBM19	RNF152	RP11-126O1.1	RP11-149B9.2	RP11-162A12.1	RP11-202D18.2	RP11-
	204C16.4	RP11-20E23.1		RP11-342D14.1		RP11-377B2.2	RP11-38O23.5	RP11-38O23.7	RP11-472N6.1
	RP11-474D14.2		RP11-479I1.4		RP11-514F8.2	RP11-524P6.1	RP11-540L11.2		RP1-154K9.2
	RP11-552E4.2		RP11-552E4.3		RP11-57G10.6	RP11-617O8.1	RP11-640A1.1		RP11-652N17.1
	RP11-702C7.2		RP11-9H16.1		RP1-22N22.1		RP1-290C9.2		RP1-50A13.1
	50A13.2		RP3-326I13.1		RP5-1174J21.1	RP5-1174J21.2	RP6-105D16.1		RP6-218J18.3
	RPL12P8	RTTN	SALL3	SCN8A	SERPINB5	SH2D4B	SIRT1	SLC38A5	SNORA68
									snoU13

SOCS6	SSX5 TNFRSF11A	TSHZ1 TXNL1	TXNL4A	U1	U2	U52112.12	U6	U7
U8	VCX3B	VENTXP1 VPS4B WDR7	Y_RNA	ZADH2	ZCCHC2 ZFX	ZNF236	ZNF407	ZNF532

PANTHER analysis: 83 mapped ids are found, 84 mapped ids are not found.

5.8.10.8 CnLOH of Triple Negative DCIS Compared to CnLOH in Triple Negative Invasive Breast Disease

There is CnLOH present in TN DCIS (n=18/19) that are not observed in TN invasive breast disease (n=0/7). This represents CnLOH for TN DCIS.

1. P-values <0.0001;
2. A gene list is mapped from 78 genomic regions found on chromosomes 1, 2, 3, 6, 7, 10, 20, 21;
3. These regions encompass 529 genes altered in TN DCIS;

Genes:

sept7	5S_rRNA	7SK	ABHD11	AC004455.1	AC004691.5	AC005154.5	AC005189.6	AC005537.2	AC005582.1
	AC006022.4	AC006466.5	AC006478.1	AC006960.7	AC006978.1	AC007392.3	AC007403.1	AC007403.2	AC009474.1
AC009474.2	AC010132.10		AC010132.11		AC010132.9	AC011738.4	AC018685.2	AC027269.2	AC063927.2
	AC063927.3	AC063927.4	AC063927.5	AC063927.6	AC063927.7	AC063927.8	AC063927.9	AC073464.11	
	AC073464.9	AC073846.1	AC073846.2	AC073846.3	AC078784.1	AC078811.1	AC087069.2	AC092265.1	AC092898.1
	AC093168.1	AC107027.2	AC108729.1	AC117477.1	AC121332.1	ADCYAP1R1	ADPRHL2	AF064859.2	AF127577.10
	AF127577.11		AF127577.12	AF127577.8	AF127936.3	AF127936.5	AF127936.6	AF127936.7	AF127936.8
AF130358.5	AF165138.6	AF165138.7	AIF1	AIM1L	AJ006998.2	AJ009632.3	AL022170.1	AL023807.1	AL031655.1
	AL031655.2	AL034372.1	AL034449.2	AL035670.1	AL049570.1	AL050322.1	AL121988.1	AL133268.1	AL136230.1
AL138724.3	AL138889.1	AL139044.1	AL139044.2	AL139286.1	AL160037.2	AL160400.1	AL160400.2	AL162579.1	AL355877.1
	AL589723.1	AL590062.1	AL591845.1	AL645811.1	ANKRD20B	ANLN	AOAH	AP000304.12	
	AP000313.1	AP000313.2	AP000318.2	AP000569.2	AP000569.8	AP001347.1	AP001347.5	AP001347.6	AP001466.1
	AP001466.7	AP001634.5	APOM	AQP1	ARHGAP31	ASS1P1	ASTE1	ATP2C1	ATP5LP2
ATP5O	ATP6V1G2	B4GALT4	BAK1		BAT1	BAT2	BAT3	BAT4	BAT5
BCL7B	BIN1	BLVRA	BPESC1	BRWD1	BZW1L1		C1orf113	C1orf212	C1orf216
C20orf187	C21orf81	C3orf16	C3orf30	C3orf72	C6orf125	C6orf115	C6orf227	C6orf25	C6orf26
									C6orf27
C6orf47	C7orf25	C7orf44	CAMK1D	CAP2	CATSPER4	CCDC21		CCHCR1	CD52
CDSN	CEP70	CHCHD6	CLDN3	CLDN4	CLIC1		CLSPN	CMAH	CNKSR1
COL21A1	COL6A6	COL8A2	COMMD2	CPNE4	CRHR2	CSF3R	CSNK2B	CST2	
CST5	CSTP1		CTD-2021J15.1		CTD-2021J15.2		CYCSP42	CYP39A1	DHAH2
									DHFRP2
DLGAP3	DNAJC30	EEF1A1P25	EIF2C1	EIF2C3	EIF2C4	ENTPD6	EPDR1	ERLEC1P1	ESYT3
EXTL1	FABP7	FAIM	FAM176B	FAM188B	FAM65B	FAM8A1	FAT1P1	FGFR3P	
FOXJ2	FTLP18	GAPDHP39	GARS	GGNBP1	GGTLA4	GHRHR	GJA4	GJB3	
GJB4	GJB5	GLI3	GMNN	GPR116	GRRP1	GS1-179L18.1		GS1-278J22.1	
GS1-278J22.2		GSTO3P	HCG27	HCP5	HECW1	HIST1H1A	HIST1H1PS2	HLA-B	HLA-
C	HLA-DQA1	HLA-DQB1	HLA-S	HMGCS2	hsa-mir-548a-1	HSF2BP	IGSF11	INMT	IP6K3

MTCO3P1	ITPR3	ITSN1	KDM1B	KIAA0319L	KIF13A	LEMD2	LIPI	LRRC16A	LSAMP	LSM10	
	LST1	LTA	LTB	LY6G5B	LY6G5C	LY6G6C		LY6G6D	LY6G6E	LY6G6F	
	MAP7D1	MCCD1	MEP1A	MICA	MICB	MLN	MLXIPL	MME	MRAS	MRPL3	
	MRPL32	MRPS15	MRPS22		MRPS24	MRPS6	MSH5	MT-ATP6	MT-CO1	MT-CO3	
	MT-ND1	MT-ND2	MT-ND3	MT-ND4	MT-ND4L	MT-ND5	MT-ND6	NCDN		NCR3	
	NCRNA00240		NEK11	NF1L1	NFKBIL1	NHLRC1	NRIP1	NUDT16	NUDT16P	NUP153	
	OR2AD1P	OR2W1	OSCP1	PA2G4P2		PABPC1P10	PAFAH2	PCBP2P1	PCCB	PDIK1L	
	PFN2	PHGDH		PIK3CB	PIK3R4	PKIB	PLA2G7	PLCB1	PLCH1	PLXNA1	
POLR2CP	POLR2LP	POM121L3P	POU5F1	PPIAP9	PRIM2	PRR23A	PRR23B		PRR23C	PSMA2	
	PSMB2	PSMG1	PSORS1C1	PSORS1C2	PSORS1C3	RAD23BLP	RBM11		RBM24	RBPM5LP	
RCAN2	RHOT1P2	RNF13	RNF144B	RP11-102J14.1		RP11-115O13.1	RP11-117F22.1			RP11-117L15.1	
	RP11-14H3.3		RP11-153N17.1		RP11-167H9.3		RP11-167H9.4			RP11-191A15.1	
	RP11-191A15.2		RP11-191A15.3		RP11-191A15.4		RP11-201E8.1			RP11-221J22.1	
	RP11-221J22.2		RP11-231P20.2		RP11-231P20.3		RP11-232M24.1			RP11-238O13.1	
	RP11-255N4.2		RP11-255N4.3		RP11-263A24.1		RP11-265F19.1			RP11-268F1.2	
	RP11-268F1.3		RP11-280F2.1		RP11-280F2.2		RP11-280H21.1			RP11-305I9.1	
	RP11-306L14.1		RP11-324H4.1		RP11-367G6.1		RP11-384F7.1			RP11-384F7.2	
	RP11-394D14.1		RP11-39E3.3		RP11-39E3.4		RP11-39E3.5			RP11-39E3.6	
	RP11-40M23.1		RP11-435D7.3		RP11-446F17.3	RP11-446H18.3		RP11-451G4.2			
	RP11-451G4.3		RP11-463H24.1		RP11-484M3.4		RP11-484M3.5			RP11-517B11.2	
	RP11-517B11.4		RP11-517B11.6		RP11-548O1.1		RP11-548O1.3			RP11-561O4.1	
	RP11-569G9.7		RP11-605F14.3		RP11-622P13.2		RP11-631B21.1			RP11-631B21.2	
	RP11-639F1.1		RP11-639F1.2		RP11-639F1.3		RP11-63E16.1			RP11-649A16.1	
	RP11-651P23.2		RP11-651P23.3		RP11-651P23.4		RP11-651P23.5			RP11-651P23.6	
	RP11-768G7.1		RP11-768G7.2		RP11-768G7.3		RP11-779P15.1			RP11-779P15.2	
	RP11-795J1.1		RP11-795J1.2	RP11-810H18.1		RP11-933H2.4		RP11-93B21.1			
	RP11-96L14.7		RP1-245M18.2		RP1-273P12.1		RP1-273P12.3			RP1-278O22.2	
	RP1-71H19.2		RP3-347E1.2		RP3-425P12.1		RP3-425P12.2			RP3-425P12.4	
RP3-425P12.5	RP3-437I16.1		RP3-468B3.2		RP3-501N12.3		RP3-522P13.1	RP3-522P13.2			
	RP3-522P13.3		RP3-525L6.2	RP4-564O4.1		RP4-635B5.1	RP4-665N4.4		RP4-665N4.5		
	RP4-697P8.2		RP4-697P8.3		RP4-718N17.1		RP4-718N17.2		RP4-728D4.2		
	RP4-728D4.3		RP4-738P15.1		RP4-738P15.4	RP4-765A10.1		RP4-765A10.2	RP4-		
796I8.1	RP5-1100I6.1		RP5-1100I6.2		RP5-877J2.1	RP5-983H21.3		RP5-997D16.2			
	RPL15P4	RPL3P2	RRP1B	SAMSN1	SCARNA17	SCARNA18	SCGN	SEPT7P3	SFRP4	SH3BGRL3	
SLC25A27	SLC30A2	SLC5A3	SMPDL3A	SNORA38	SNORA58	SNORA63	SNORA73	SNORD45		SNORD83	
	snoU13	STAG1	STK40	STX1A	TBL2	TCF19	TDRD6	TEKT2	TFAP2E	THRAP3	
	TMEM45A	TNF	TPMT	TRAPPC3	TRIM38	TRIM63	TXNDC3	U1	U3	U4	U5
U6	U7	UBXN11	UPK1B	UQCRHP1	URGCP	USP25	USP8P	VAR5		VN1R11P	
	VN1R13P	VPS37D	WASF5P	WBSCR22	WBSCR26	WBSCR27	WBSCR28	WWTR1		XXbac-	
BPG181B23.4		XXbac-BPG181B23.6	XXbac-BPG248L24.10		XXbac-BPG248L24.11		XXbac-BPG248L24.12		XXbac-		
BPG254F23.5		XXbac-BPG254F23.6	XXbac-BPG254F23.7		XXbac-BPG296P20.14				XXbac-BPG296P20.15		
		XXbac-BPG299F13.14	XXbac-BPG32J3.18		XXbac-BPG3 J3.19		XXyac-YX60D10.1		Y_RNA		
	Z93017.1	ZDHHC20P2	ZMYM4	ZNF593	ZNF683	ZPLD1					

PANTHER analysis: 208 mapped ids are found, 321 mapped ids are not found.

There is CnLOH present in TN invasive breast disease (n=6/7) that are not observed in TN DCIS (n=0/19).

1. p-values < 0.05;
2. A gene list is mapped from 40 genomic regions found on chromosomes 1, 2, 3, 10, 11, 12, 19, 20;
3. These regions encompass 291 genes altered in TN invasive breast disease;

Genes:

ZSWIM4 5S_rRNA 7SK ABRA AC004016.1 AC004016.2 AC005082.1 AC005082.12 AC005082.2 AC006039.1 AC006039.4 AC006039.5
AC008686.2 AC009908.1 AC011118.2 AC011626.1 AC016405.1 AC017084.1 AC017084.6 AC021876.3 AC021876.4 AC022021.2
AC022201.4 AC023590.1 AC023632.1 AC025370.1 AC025370.2 AC025647.3 AC027238.1 AC027419.1 AC048346.1 AC064807.1
AC068228.4 AC068228.5 AC084083.1 AC084083.2 AC084114.1 AC090192.1 AC090192.2 AC090193.2 AC090802.1 AC090811.1 AC090921.
AC090922.1 AC099680.1 AC103816.1 AC104212.2 AC120053.1 ADD2 AL049651.1 AL136985.2 AL160287.2 AL161932.2 AL451107.1
AL512654.1 ANGPT1 ANKRD46 ANXA13 AP000428.1 AP002852.1 AP002907.1 AP003550.10 APBB1IP ARHGAP12 ATAD2
ATP6V1G1P4 BAALC BRD7P6 C10orf50 C7orf30 C7orf31 C8orf37 C8orf39 C8orf54 C8orf56 C8orf76
C8orf85 CACHD1 CCND1 CCNY CDH17 CLSTN2 COLEC10 CREM CSDA CSGALNACT2
CSTF2T CTC-788C1.1 CTHRC1 CUL2 CYB5P4 CYCS DCAF13 DEPDC1 DERL1 DFNA5 EXT1
EYA1 FABP5 FAM159A FAM83A FAM91A1 FAM92A1 FBXO32 FBXO43 FER1L6 FGF19
FXSD4 FZD6 GAPDP2 GEM GLUDP5 GPNMB GPX7 HAS2 HAS2AS HNRNPF hsa-mir-1273hsa-
mir-23a hsa-mir-24-2 hsa-mir-27a hsa-mir-548d-1 IAPP IGF2BP3 KB-1137H10.1 KB-1247B1.1KB-1554H10.1
KB-1615E4.1KB-1980E6.1KIAA0196 KIF5B KLF10 KLHL38 KLHL7 KRT18P3 MACROD1 MAL2 MED30
MPP7 MTSS1 MYSM1 NCALD NDUFB9 NIPAL2 NOV NRP1 NSMCE2 NUPL2 ODF1
ORAOV1 OSBPL3 OVCH1 OXR1 PABPC1 PCDH15 PCMTD1 PDP1 PDSS1 PDZK1IP1 PIK3C2G
POP1 PRB3 PRH1 PRH2 PRKG1 PRR4 RAD54B RASGEF1A RAVR2 RBM12B RET RGS22
RNF139 RNF19A RP11-1082L8.1 RP11-1110J8.1RP11-128B16.3 RP11-13L2.1 RP11-13L2.2RP11-14C22.3
RP11-14C22.4 RP11-14C22.6 RP11-152P17.1 RP11-167O6.2 RP11-168L22.2 RP11-168O22.1 RP11-
189G24.2 RP11-200A13.1 RP11-200A13.2RP11-241I20.1 RP11-241I20.3 RP11-241I20.4 RP11-241I20.5
RP11-273P3.1 RP11-297A16.2 RP11-30P9.1 RP11-318G8.1 RP11-324I22.2 RP11-342D11.2
RP11-351M16.1 RP11-351M16.2 RP11-351M16.3 RP11-359G22.2 RP11-414D17.1 RP11-472N13.3 RP11-484J3.1 RP11-51B10.3 RP11-
51B10.4 RP11-539E17.1 RP11-668N23.1 RP11-68L1.1RP11-68L1.2RP11-778D12.1 RP1-18D14.2 RP1-18D14.3 RP1-
18D14.4 RP11-941H19.1 RP11-941H19.2 RP11-959I15.1 RP13-16H11.1 RP13-16H11.2 RP13-16H11.5
RP13-16H11.6 RP13-16H11.7 RP3-388N13.1 RP4-694A7.2 RP4-694A7.3 RP4-694A7.4
RP4-753D10.3 RP4-753D10.5 RP4-794H19.2 RP5-1033K19.2 RPE65 RRM2B RSP02
SAMD12 SLC25A32 SLC30A8 SLC01A2 SLC01B1 SNORA20 SNORA31 SNORA32 SNORD77 snoU13 SNX31
SPAG1 SQLE SSTR4 STYK1 TAS2R10 TAS2R13 TAS2R14 TAS2R19 TAS2R20 TAS2R31
TAS2R42 TAS2R50 TAS2R7 TAS2R8 TAS2R9 TASP1 TATDN1 TCEB1P18 TGFA THBD TMEM65
TMEM67 TMEM74 TMTC1 TNFRSF11B TPD52 TRHR TRIB1 TRIM42 TRMT12 TRPS1 U3 U4
U6 U6atac U7 UBR5 WAC WDR67 WDYHV1 Y_RNA ZFPM2 ZHX1 ZHX2 ZNF572

PANTHER analysis: 140 mapped ids are found, 151 mapped ids are not found.

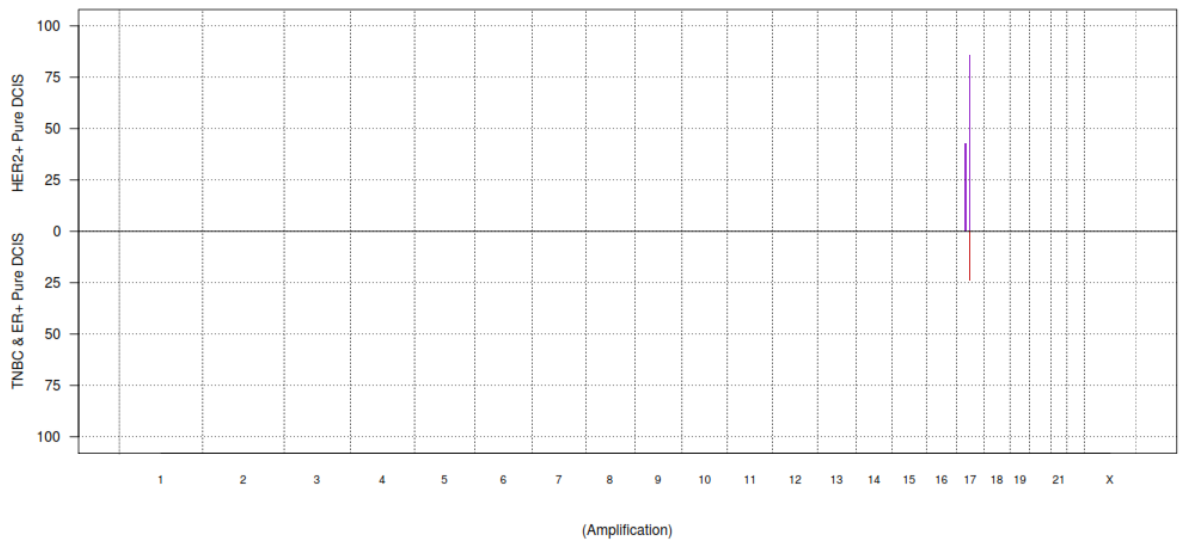
5.8.11 Copy Number Aberrations for HER2 Positive DCIS Compared to Oestrogen Receptor Positive DCIS and Triple Negative DCIS.

These series examines the difference between Her2 positive pure DCIS and oestrogen receptor positive DCIS and triple negative positive DCIS

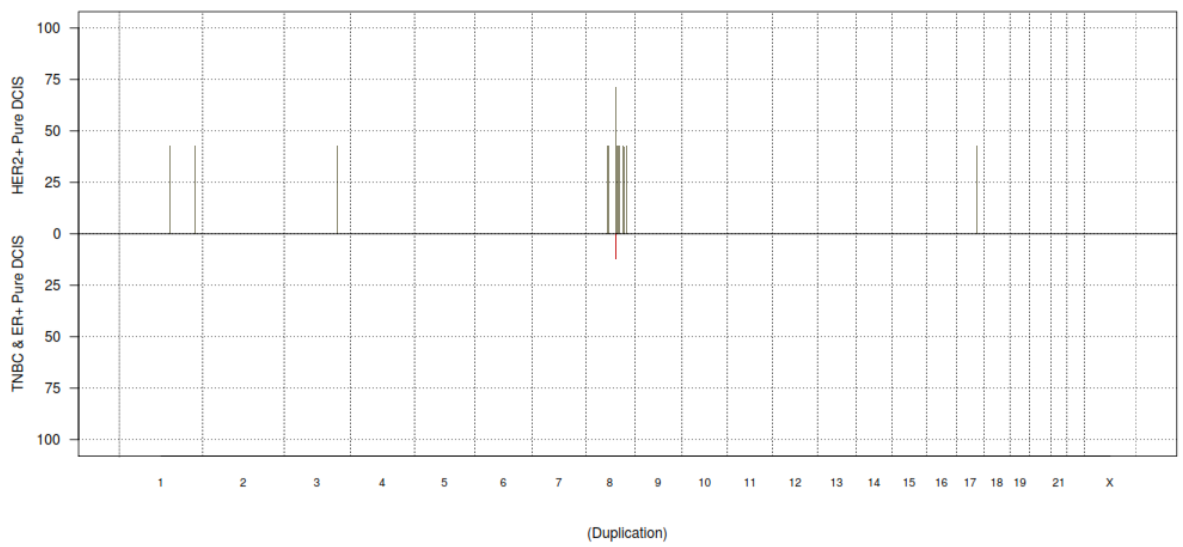
Frequency plots showing copy number aberrations between HER2 positive DCIS compared to oestrogen receptor positive DCIS and triple negative DCIS were provided by Breakthrough Breast Cancer/Research Oncology, King's College London Bioinformatics Department (Figure 44).

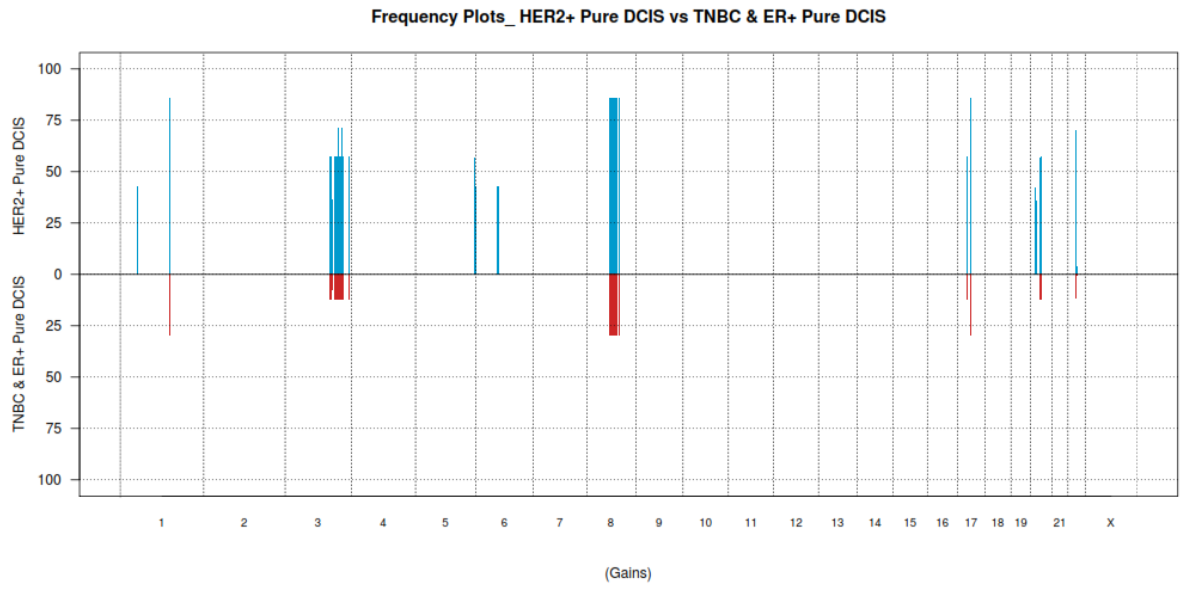
Figure 44: Frequency plots showing copy number aberrations between HER2 positive pure DCIS compared to oestrogen positive dcis and triple negative DCIS (amplifications, duplications, gains, Sc gains, losses, total losses CdLOH and CnLOH,) (pages 279-282).

Frequency Plots_ HER2+ Pure DCIS vs TNBC & ER+ Pure DCIS

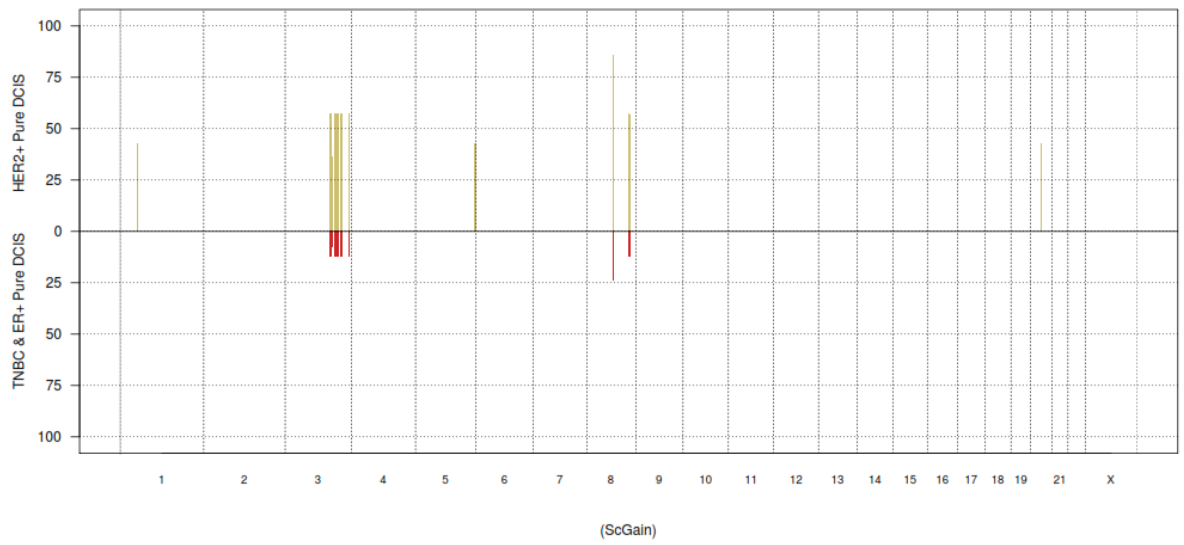


Frequency Plots_ HER2+ Pure DCIS vs TNBC & ER+ Pure DCIS

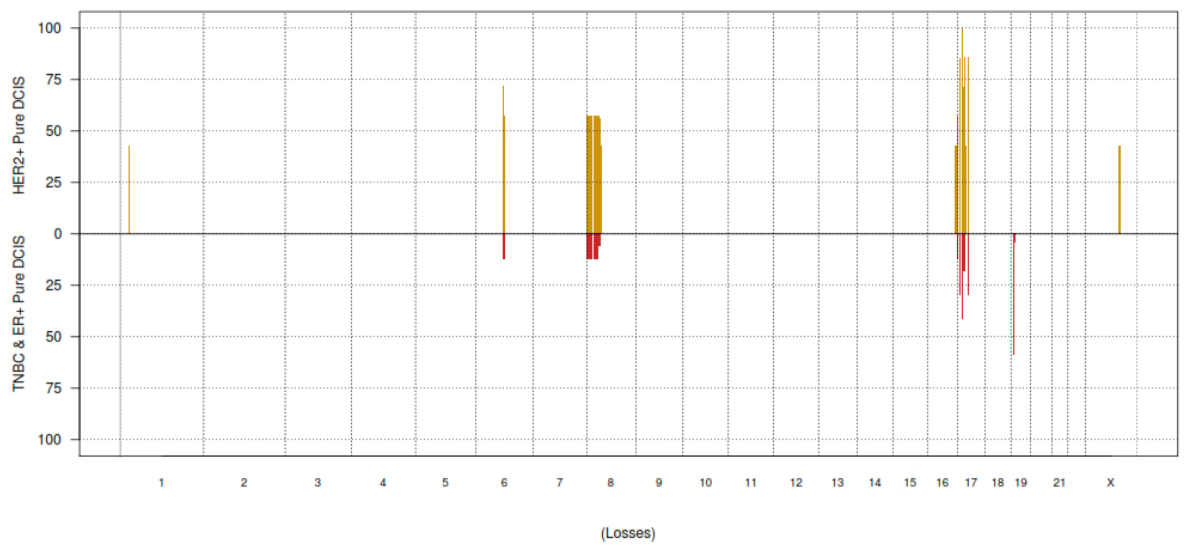


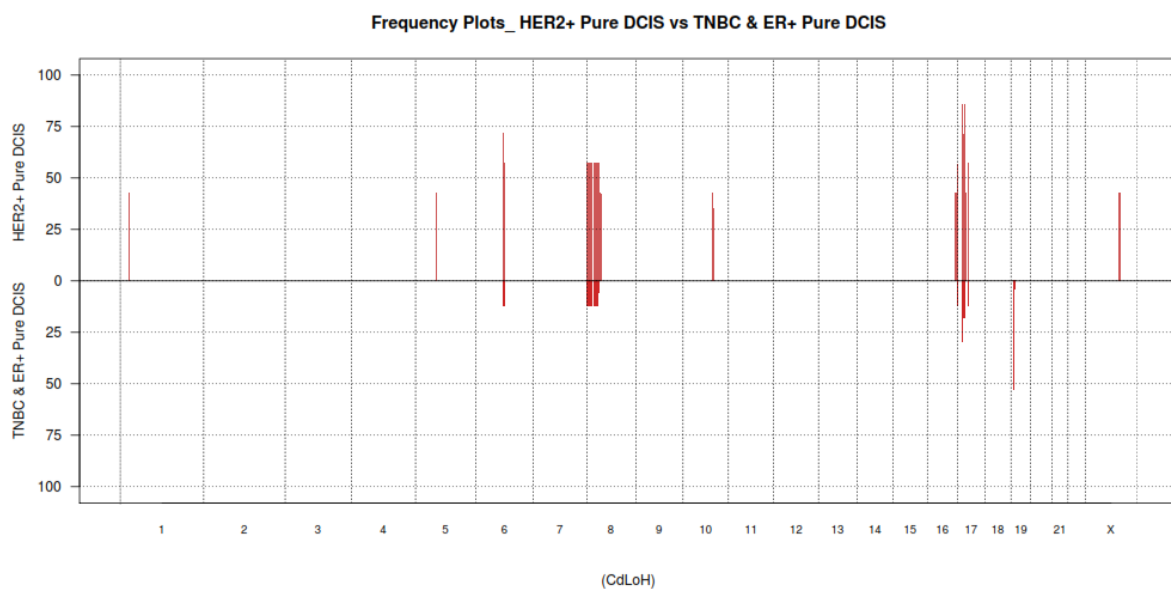
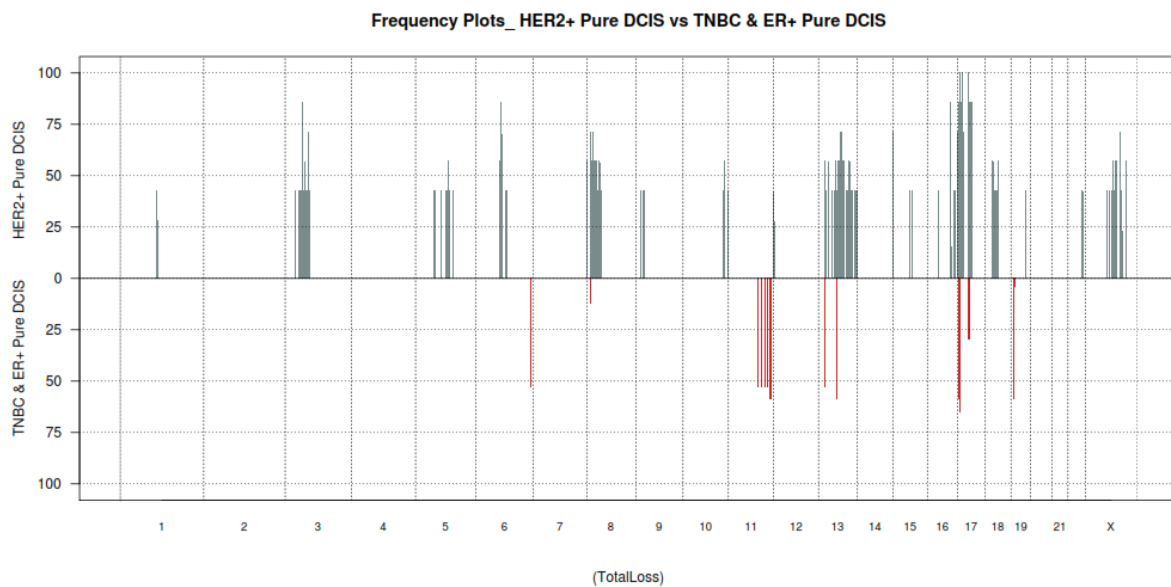


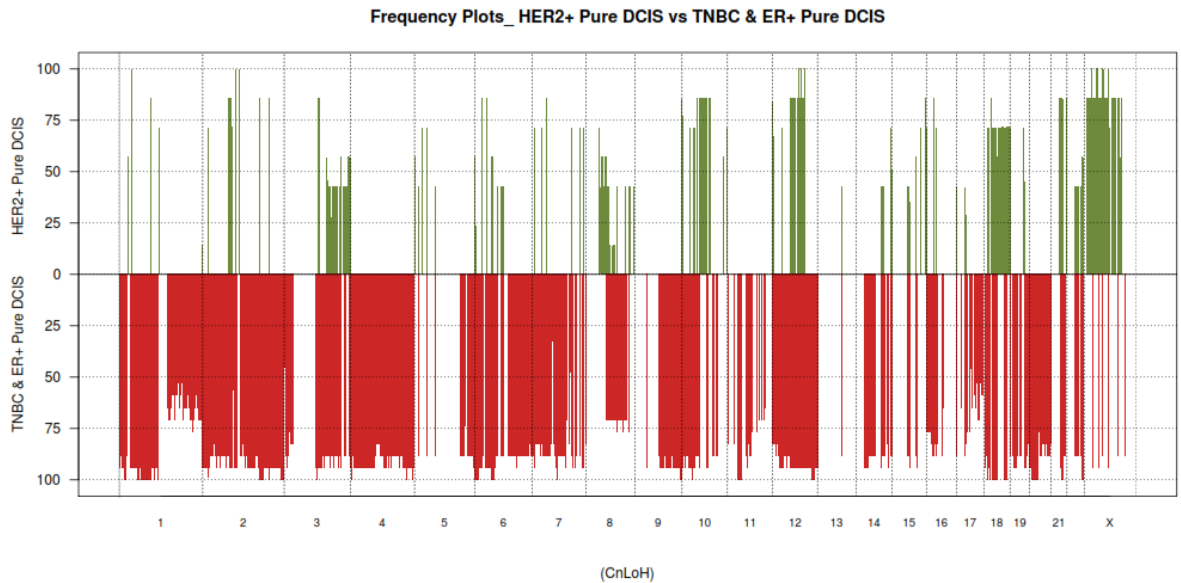
Frequency Plots_ HER2+ Pure DCIS vs TNBC & ER+ Pure DCIS



Frequency Plots_ HER2+ Pure DCIS vs TNBC & ER+ Pure DCIS







5.8.11.1 Amplification in HER2 Positive DCIS Compared to ER Positive DCIS and Triple Negative DCIS.

There are overlaps between the amplifications found in HER2 positive pure DCIS (n=6/7) and all other pure DCIS (TN and ER n=4/17).

1. p-values < 0.05;
2. A gene list is mapped from 7 regions on chromosome 17;
3. These regions encompass 45 genes altered in HER2 positive DCIS;

Genes:

AC002094.1	AC002094.2	AC002094.3	AC002094.4	AC002094.5	AC015917.1	AC015917.2	AC015917.3	AC061975.1	AC061975.2
AC061975.3	AC061975.4	AC061975.5	AC061975.6	AC079199.1	AC079199.2	AC087491.2	C17orf37	CRKRS	EFCAB5
ERBB2	GRB7	IFT20	IKZF3	NEUROD2	PGAP3	PNMT	POLDIP2	PPP1R1B	PPY2
RP11-338L22.1	SARM1	SLC13A2	SLC46A1	snoU13	SSH2	STARD3	TCAP	TMEM199	
TMEM97	TNFAIP1	U6	VTN	ZPBP2					

PANTHER analysis: 20 mapped ids are found, 25 mapped ids are not found.

5.8.11.2 Duplication of HER2 Positive DCIS Compared to Oestrogen Positive DCIS and Triple Negative DCIS.

There were duplications present in HER2 positive DCIS (n=3/7) not observed in ER and TN pure DCIS (n=0/17).

1. p-values < 0.05;
2. A gene list is mapped from 14 regions found on chromosomes 1, 3, 8, 17;
3. These regions encompass 91 genes altered in HER2 positive DCIS;

Genes:

5S_rRNA	AC005746.1	AC005901.1	AC011118.2	AC083928.1	AC092811.1	AC104958.1	AC120042.2	AC120053.1	AP003357.3	
ARFGEF1	ARL14	BCAS3	C17orf82	C1orf46	C1orf68	C8orf39	C8orf44	C8orf46	COL14A1	COPS5
CSMD3	CSPP1	CTC-820M8.1	DEPDC6	DNAH14	ENAH		FAM92A1	GGH	IVL	KB-1589B1.1KB-
1683C8.1	KPRP	LAPT4M4B	LBR	LCE1A	LCE1B	LCE1C	LCE1D	LCE1E	LCE1F	LCE2A
	LCE2B	LCE2C	LCE2D	LCE3A	LCE3B	LCE3C	LCE3D	LCE3E	LCE4A	LCE6A
	LRRC67	MATN2	MRPL13	MTBP	MTDH	MYBL1	PDP1	RBM12B	RP11-145A3.1	
	RP11-145A3.2	RP11-145A3.4		RP1-13P20.6		RP11-453N18.1		RP11-496N12.6		RP1-20N18.4
RP1-43O17.1		RP1-43O17.2		RP1-52J10.9	RP3-388N13.1		SMCP	SNORA31	SNTB1	
SPRR1A	SPRR1B	SPRR2A	SPRR2B	SPRR2D		SPRR2E	SPRR3	SPRR4	SRP9	TBX2
TBX4	TMEM67	TTPA	U6	U7	VCPIP1	Y_R				

PANTHER analysis: 58 mapped ids are found, 33 mapped ids are not found.

5.8.11.3 Genomic Gains in HER2 Positive DCIS Compared to Oestrogen Positive DCIS and Triple Negative DCIS.

There are genomic gains in HER2 positive pure DCIS (n=4/7) not present in ER and TN pure DCIS (n=0/17).

1. p-values < 0.05;
2. A gene list is mapped from 10 genomic regions found on chromosomes 1, 5, 6, 20;
3. These regions encompass 92 genes altered in HER2 positive pure DCIS;

Genes:

AACSL	AC022096.2		AC027317.2	AC027317.2	AC091934.1	AC104115.1	AC104117.1	AC104117.2	AC136940.3	AC145098.1
AC145098.2	AC145098.3	AC146507.1	ADAMTS2	AGXT2L2	AL050329.1	AL589736.1	CNOT6	COL23A1	CTB-129O4.1	
CTC-573N18.1		CTD-2301A4.1		CTD-2301A4.3		DBN1	DDX41	DMRTA2	DOK3	
EIF4E1B	EYS	F12	FAF1	FAM193B	FGFR4	FGFR4	FKBP1C	FLT4	GCNT1P4	
GFPT2	GRK6	GRM6	HNRNPAB	LGSN	LMAN2	MACROD2	MAPK9	MXD3	NSD1	

OR2AI1P	OR2Y1	PDLIM7	PFN3	PHF3	PRELID1	PRR7	PTP4A1	RAB24	RASGEF1C
RGS14	RP11-1334A24.4		RP11-1334A24.5		RP11-1334A24.6		RP11-164N20.1		RP11-183G22.1
RP11-183G22.2		RP11-183G22.3		RP11-184C23.1		RP11-184C23.3		RP11-184C23.4	
RP11-252I14.1		RP11-349P19.1		RP11-442B12.1		RP11-448N11.2		RP11-451H23.1	
RP11-451H23.2		RP11-451H23.3		RP11-59D5__B.2		RP11-59D5__B.3		RP3-407E4.2	RP3-
407E4.3	RP3-407E4.4		RP5-1148A21.1	SCGB3A1	SLC34A1	SNCB	TSPAN17	U1	U1
U6	ZFP2	ZNF346	ZNF346	ZNF354A	ZNF354B	ZNF354C	ZNF454	ZNF879	

PANTHER analysis: 47 mapped ids are found, 45 mapped ids are not found.

There are genomic gains in TN and ER pure DCIS (n=5/17) with associated genomic gains in HER2 positive pure DCIS (n=6/7).

1. p-values < 0.05;
2. A gene list is mapped from 14 genomic regions found on chromosomes 1,8,17.
3. These regions encompass 129 genes altered in TN v DCIS, ER pure DCIS and HER2 positive pure DCIS.

Genes:

5S_rRNA	AC012533.1	AC016057.1	AC016240.1	AC022598.1	AC022861.1	AC022861.2	AC022861.3	AC022861.4	AC022861.5	
	AC022861.6	AC022861.7	AC023644.1	AC024367.1	AC027006.1	AC037450.1	AC083928.1	AC087491.2	AC091175.1	AC099805.1
AC103816.1	AC104012.1	AC104212.2	AC110998.1	AC132219.1	AC132219.2	AF121898.1	AL592492.2	ARFGEF1	C8orf44	
C8orf45	C8orf46	CHMP4C	CNBD1	COPS5	CPA6	CSPP1	DCAF4L2	E2F5	ERBB2	
EYA1	FABP12	FABP4	FABP5	FABP9	FAM164A	HEY1	IKZF3	IL7	IMPA1	
JPH1	LRRRC67	LRRCC1	LY96	MMP16	MRPS28	MSC	MYBL1	NBPF10	NBPF14	
NBPF15	NBPF16	NEUROD2	PAG1	PGAP3	PKIA	PMP2	PNMT	PPIAL4A	PPP1R1B	
PTTG3P	RALYL	RP11-289I10.1		RP11-289I10.2		RP11-34M16.1		RP11-34M16.2		
RP11-367C12.1		RP11-367E12.1		RP11-453N18.1		RP11-48B3.1		RP11-495P10.1		
RP11-495P10.2		RP11-508K19.1		RP11-656G20.1		RP11-666A1.1		RP11-666A1.2		
RP11-666A1.3		RP11-666A1.4		RP11-666A1.5		RP11-69I13.1		RP11-6I2.1	RP11-763B22.2	
RP11-763B22.3		RP11-763B22.4		RP11-7F18.1	RP11-89F3.2	RP11-89F3.3	RP11-91G11.1		RP11-941H19.1	
RP11-941H19.2		RP11-98H4.1		RP6-74O6.1	SGK3	SLC10A5	SNHG6	SNORA20	SNORD87	
snoU13	SNX16	STARD3	STMN2	TCAP	TCEB1	TMEM70	TPD52	TRPA1	U1	U11
U2	U6	U7	U8	VCPIP1	WWP1	Y_RNA	ZBTB10	ZFAND1	ZFHX4	
ZNF704										

PANTHER analysis: 56 mapped ids are found, 73 mapped ids are not found.

5.8.11.4 Genomic Sc Gains in HER2 Positive DCIS Compared to Oestrogen Positive DCIS and Triple Negative DCIS.

There were Sc gains present in HER2 positive pure DCIS (n=6/7) and ER and TN pure DCIS (n=4/17).

1. p-values < 0.05;
2. A gene list is mapped from 10 regions found on chromosomes 1, 3, 5, 8, 20;
3. These regions encompass 188 genes altered in HER2 positive pure DCIS and ER and TN pure DCIS;

Genes:

5S_rRNA	7SK	A4GNT	AACSL	AADAC	AADACL2	AC022096.2	AC026320.2	AC027317.2	AC087667.1	
AC091934.1	AC092965.1	AC104117.1	AC104117.2	AC104629.1	AC104629.2	AC104629.3	AC104629.4	ADAMTS2	ARMC8	
C3orf36	C3orf55	C3orf72	CCDC50	CCRL1	CDV3	CEP63	CEP70	CLDN18	DBR1	DEFB121
DEFB122	DEFB123	DEFB124	DHX36	DMRTA2	DNAJC13	DZIP1L	EEF1A1P25		EIF4E1B	EPHB1
ESYT3	FAF1	FAIM	FGF12	FGFR4	FOXL2	GAPDHP39	GFM1	GPR149	GPR79	GRM6
HM13	IL20RB	IQCJ	KY	LXN	MFS1	MLF1	MRAS	MSL2	NCK1	
NCRNA00028	NPHP3	OSTN	OTTHUMG00000159780	PCCB	PDCD10	PHC3	PIK3CB	PPP2R3A		
PRKCI	PYDC2	RAB6B	RARRES1	REM1	RP11-102M11.1		RP11-113A11.1		RP11-117F22.1	
RP11-12N13.1		RP11-12N13.2		RP11-12N13.3	RP11-12N13.4		RP11-12N13.5		RP11-12N13.6	
RP11-167H9.1		RP11-183G22.1		RP11-183G22.2		RP11-183G22.3		RP11-192K2.3		
RP11-197K6.1		RP11-202A13.1		RP11-202A13.2		RP11-206M11.4		RP11-206M11.7		RP11-
217E22.1	RP11-217E22.2		RP11-217E22.5		RP11-237P21.1		RP11-23D24.1		RP11-23D24.2	
RP11-278L15.1		RP11-278L15.2		RP11-290K4.1		RP11-290K4.2		RP11-293N1.1		
RP11-2A4.1	RP11-2A4.3	RP11-2G17.1	RP11-305O4.1		RP11-305O4.2		RP11-333H9.6		RP11-379F4.1	
RP11-379F4.4		RP11-404G16.1		RP11-404G16.2		RP11-413G22.2		RP11-440K22.1		
RP11-450H5.1		RP11-452B4.1		RP11-454C18.1		RP11-454C18.2	RP11-463H24.1		RP11-463H24.4	
RP11-468F18.2		RP11-538P18.1		RP11-538P18.2		RP11-548O1.1		RP11-548O1.3		
RP11-575C1.1		RP11-576M8.2		RP11-598P12.1		RP11-642L11.1		RP11-64D22.1		
RP11-64D22.2	RP11-64D22.5		RP11-655G22.1		RP11-655G22.2		RP11-656G9.3		RP11-657O9.1	
RP11-65E22.3		RP11-731C17.1		RP11-785G3.1		RP11-79L9.2	RP11-79M21.2		RP11-79M21.3	
RP11-85F14.1		RP11-85F14.5		RP11-85F14.6		RP11-91K8.2		RP13-81O8.1		RP3-
324O17.4	RSRC1	RYK	SERPINI1	SERPINI2	SHOX2	SLCO2A1	SNCB	SNORA33	snoU13	SOX14
SRPRB	STAG1	SUCNR1	TF	TM4SF1	TM4SF18	TMEM22	TOPBP1	TSPAN17	TXNDC6	U1
U6	U7	U8	UBA5	UTS2D	VEPH1	WDR49	Y_RNA	ZBBX	ZFP2	ZNF346
ZNF354B	ZNF354C	ZNF454	ZNF879							

PANTHER analysis: 82 mapped ids are found, 106 mapped ids are not found.

5.8.11.5 Losses in HER2 Positive DCIS Compared to Oestrogen Positive DCIS and Triple Negative DCIS.

There were losses present in HER2 positive pure DCIS (n=3/7) not observed in ER and TN positive pure DCIS (n=0/17).

1. p-values < 0.05;
2. A gene list is mapped from 9 regions found on chromosomes 1, 8, 15, 16, X;
3. These regions encompass 87 genes altered in HER2 positive pure DCIS;

Genes:

AC009113.1	AC009113.2	AC009113.3	AC010531.1	AC010531.2	AC027702.2	AC092139.1	AC092139.2	AC092384.1	AC092384.2
AC099524.1	AC116552.1	AC136285.1	AC138028.1	AC138028.2	AC138028.3	ACSF3	AGPAT6	AHDC1	AL049610.1
ANK1	AP3M2	APRT	C16orf81	C16orf85	CBFA2T3	CD164L2	CDH15	CDT1	CTU2
CYBA	FAM38A	FBXO31	FCN3	GALNS	GAN	GIN54GLRA4	GOLGA7	GPR3	hsa-mir-486
IKBKB	IL17C	LL0XNC01-250H12.2	LL0XNC01-250H12.3	MAP1LC3B	MAP3K6	MORF4L2	MVD	MYST3	
NGFRAP1	NKX6-3	PABPN1L	PLAT	PLCG2	PLP1	RAB40A	RAB9B	RNF166	RP11-173M1.1
RP11-178L8.1		RP1-144C9.2		RP11-589C21.1		RP1-159A19.4		RP4-752I6.1	RP5-1055C14.4
RP5-1055C14.6		RP5-1055C14.7		RP5-1142A6.2		SLC20A2	SNAI3	SNORD112	snoU13
snoU13	SYTL1	TCEAL1	TCEAL3	TCEAL4	TMEM31	TRAPPC2L	WASF2	Y_RNA	Z73964.1
ZC3H18	ZCCHC14	ZFPM1	ZNF469	ZNF778					

PANTHER analysis: 50 mapped ids are found, 37 mapped ids are not found.

There were losses present in ER and TN pure DCIS (n=10/17) not observed in any HER2 positive pure DCIS (n=0/24).

1. P-values < 0.05;
2. A gene list is mapped from 3 genomic regions from chromosome 19;
3. These regions encompass 3 genes altered in non HER2 DCIS;

Genes:

MUC16	AC008734.1	AC008734.2
-------	------------	------------

PANTHER analysis: 1 mapped ID found, 2 mapped ids are not found.

5.8.11.6 Total Loss in HER2 Positive DCIS Compared to Oestrogen Positive DCIS and Triple Negative DCIS.

There were total losses present in HER2 positive pure DCIS (n= 7/7) also present in some ER and TN pure DCIS (n=4 /17).

1. p-values < 0.05;
2. A gene list is mapped from 28 regions found on chromosomes 3, 6, 16, 17;
3. These regions encompass 293 genes altered in HER2 positive pure DCIS and ER and TN pure;

Genes:

5S_rRNA	7SK	AARS	AC000003.1	AC003664.1	AC003958.1	AC003958.2	AC003958.5	AC003958.6	AC004223.2
AC004675.1	AC005224.2	AC005277.1	AC005304.1	AC005304.2	AC005304.3	AC005358.1	AC005358.3	AC005410.2	AC006070.11
AC006070.12		AC006070.13		AC006070.7	AC009060.1	AC009060.2	AC009060.3	AC011193.1	AC012184.1
AC015842.1	AC015849.1	AC015849.4	AC019349.4	AC024610.1	AC025335.1	AC025518.1	AC026468.1	AC026954.6	AC027045.1
AC069363.1	AC087742.1	AC091178.1	AC100793.1	AC100808.10		AC100808.7	AC104581.1	AC107993.1	AC113189.5
AC123769.1	AC126327.1	AC126327.4	AC129492.6	AC131056.1	AC135178.7	ACADVL	ACAP1	ACCN1	ALOX12B
ALOX15B	ALOXE3	AMAC1L3	ANKFY1	ARHGEF15	ARL5C	ATP6V0A1	AURKB	BECN1	C17orf102
C17orf44	C17orf59	C17orf61	C17orf68	C17orf74	C17orf81	C17orf96	C17orf98	CACNB1	CCDC56
CCL1	CCL11	CCL13	CCL14	CCL15	CCL16	CCL18	CCL2	CCL23	CCL3
CCL3L1	CCL4	CCL4L1	CCL4L2	CCL7	CCL8	CCR10	CCT6B	CD68	CDRT15
CDRT15P	CHD3	CHRN1	CISD3	CLDN7	CLEC18A	CLEC18C	CNTD1	CNTNAP1	CNTROB
COASY	COG4	COX10	CTB-75G16.1		CTD-2303K11.1		CTD-3193K9.1		CWC25
CYB5D1	CYB5D2	DDX19A	DDX19B	DHRS11	DHRS7C	DLG4	DNAH2	DNAH9	DULLARD
DUSP7	DVL2	EFNB3	EIF4A1	EIF5A	ELAC2	EXOSC6	EZH1	FAM134C	FBXO47
FGF11	FNDC8	FUK	FXR2	GABARAP	GAS7	GGNBP2	GPS2	GUCY2D	HES7
HS3ST3A1	hsa-mir-140	hsa-mir-1972	hsa-mir-324	hsa-mir-548h-3		hsa-mir-744	HSD17B1	HSD17B1P1	KCNAB3
KCTD11	KDM6B	KRBA2	KRT31	KRT33A	KRT33B	KRT34	KRT37	KRT38	KRTAP16-1
KRTAP17-1	KRTAP2-1	KRTAP2-4	KRTAP29-1	KRTAP4-1	KRTAP4-12	KRTAP4-14	KRTAP4-2	KRTAP4-3	KRTAP4-4
KRTAP4-5	KRTAP4-6	KRTAP4-7	KRTAP4-8	KRTAP4-9	KRTAP9-1	KRTAP9-2	KRTAP9-3	KRTAP9-4	KRTAP9-6
KRTAP9-7	KRTAP9-8	KRTAP9-9	LASP1	LIG3	LSMD1	LYZL6	MAP2K4	MLLT6	MLX
MPDU1	MRM1	MYH10	MYO19	MYOCD	NAGLU	NDEL1	NEURL4	NLE1	NLGN2
ODF4	PCGF2	PDXDC2	PER1	PFAS	PHF23	PIGW	PIP5K2B	PLEKHH3	PLSCR3
PLXDC1	POLR2A	PSMB3	PSMC3IP	PSME3	PTRF	RAD51L3	RAMP2	RANGRF	RDM1
RFFL	RNF222	RP11-106J23.1		RP11-1096G20.1		RP11-462C21.1		RP11-554D15.1	
RP11-554D15.3		RP11-554D15.4		RP11-565F19.1		RP11-713H12.1		RP5-837J11.1	RPL19
RPL23	RPL26	RPS27P26	SCARNA21	SEN3	SLC25A35	SLC2A4	SNORA21	SNORA48	SNORA67
SNORA69	SNORA74	snoU13	snoU6-77	SOX15	SPEM1	ST3GAL2	STAC2	STAT3	STX8
TBC1D3B	TBC1D3C	TBC1D3G	TBC1D3H	TMEM102	TMEM107	TMEM132E	TMEM88	TMEM95	TNFSF12
TNFSF12-TNFSF13		TNFSF13	TNK1	TP53	TRAPP1	TUBG1	TUBG2	U11	U6
U8	UBE2G1	USP43	VAMP2	Vault	VPS25	WDR16	WDR51A	WNK4	WRAP53
WWP2	Y_RNA	YBX2	ZBTB4	ZNF18	ZNF830	ZNHIT3	ZZEF1		

PANTHER analysis: 192 mapped ids are found, 101 mapped ids are not found.

There were total losses present in more than half of the ER pure DCIS and TN pure DCIS (n=11/17) not present in any of the HER2 positive pure DCIS cases (n=0/7)

1. p-values < 0.05;
2. A gene list is mapped from 42 genomic regions found on chromosomes 6,11,13,17, 19.
3. These regions encompass 67 genes altered in ER pure DCIS and TN pure DCIS.

Genes:

5S_rRNA	AC008734.1	AC008734.2	AC087498.1	AC090282.1	AC097370.1	ACRV1	ACTL9	ADAMTS10	AL136359.1
AL137001.1	AP001482.1	AP003027.2	ASPA	C11orf87	CEP164	CHEK1	CTD-2557P19.1	DDX25	
GRAMD1B	GRM5	hsa-mir-1253		HYLS1	MUC16	OLFM4	OR10D1P	OR10D3P	OR1A1
OR1D2	OR1D5	OR1E1	OR1E2	OR1G1	OR2Z1	OR3A1	OR3A2	OR3A3	OR3A4
OR8D1	OR8D2	OR8F1P	OR8G2P	OR8G5	PARK2	PATE1	PATE2	PATE3	PATE4
									PUS3
RAP1GAP2	RP11-24H2.1		RP11-24H2.2		RP11-301J16.2		RP11-301J16.3		RP11-301J16.5
RP11-301J16.7	RP11-442J17.4		SCN3B	SPATA22	TRPV3	U1	U6	VWA5A	WSCD1
ZNF202									

Genes PANTHER analysis: 39 mapped ids are found, 28 mapped ids are not found.

5.8.11.7 CdLOH in HER2 Positive DCIS Compared to Oestrogen Positive DCIS and Triple Negative DCIS.

There is CdLOH present in HER2 positive pure DCIS (n= 4/7) not observed in any of the ER and TN pure DCIS (n= 0/17).

1. p-values < 0.05;
2. A gene list is mapped from 13 regions found on chromosomes 1, 5, 8, 10, 16, 17, X;
3. These regions encompass 83 genes altered in HER2 positive pure DCIS;

Genes:

AC009113.1	AC009113.2	AC009113.3	AC010531.1	AC010531.2	AC092139.1	AC092139.2	AC092384.1	AC092384.2	AC099524.1
AC116552.1	AC136285.1	AC138028.1	AC138028.2	AC138028.3	ACSF3	AGPAT6	AHDC1	AL049610.1	ANK1
APRT	ATAD1	C16orf81	C16orf85	CBFA2T3	CD164L2	CDH15	CDT1	CFLP1	CTU2
CYBA	FAM38A	FBXO31	FCN3	GALNS	GAN	GINS4	GLRA4	GOLGA7	GPR3
									hsa-mir-486
IKBKB	IL17C	LL0XNC01-250H12.2	LL0XNC01-250H12.3			MAP1LC3B	MAP3K6	MORF4L2	MVD
									NGFRAP1
NKX6-3	PABPN1L	PAPSS2	PLAT	PLCG2	PLP1	RAB40A	RAB9B	RNF166	RP11-173M1.1
RP11-178L8.1		RP1-144C9.2		RP1-159A19.4		RP4-752I6.1	RP5-1055C14.4		RP5-1055C14.6

RP5-1055C14.7	RP5-1142A6.2	SLC20A2	SNAI3	snoU13	SYTL1	TCEAL1	TCEAL3	
TCEAL4	TMEM31	TRAPPC2L	WASF2	Z73964.1	ZC3H18	ZCCHC14	ZFPM1	ZNF469

PANTHER analysis: 50 mapped ids are found, 33 mapped ids are not found.

There is CdLOH present in ER pure DCIS (n= 9/17) not observed in HER2 Positive pure DCIS (n=0/7)

1. p-values < 0.05;
2. A gene list is mapped from 1 genomic regions found on chromosomes 19;
3. These regions encompass 1 gene altered in ER pure DCIS;

Gene: Muc16

PANTHER analysis: 1 mapped id is found.

5.8.11.8 CnLOH in HER2 Positive DCIS Compared to Oestrogen Positive DCIS and Triple Negative DCIS.

There is CnLOH present in HER2 positive pure DCIS (n=7/7) not observed in ER and TN pure DCIS (n=0/17)

1. p-values < 0.05;
2. A gene list is mapped from 14 regions found on chromosomes 1, 2, 12, X;
3. These regions encompass 104 genes altered in HER2 positive pure DCIS;

4. **Genes:** AC117494.15S_rRNA AC004074.1 AC004074.3 AC004074.4 AC004673.1 AC092198.1 AC117517.1 AF241726.2
AF241726.4 AF241726.6 AL031115.1 AL121578.2 AL161779.1 AL591845.1 AL627402.1 ATP6AP2 BCOR
C1orf113 C2orf55 CASK CTD-2324A24.1 CXorf38 CYBB DDX3X DDX53 DYNLT3
FAM176B FAM3C2 FTLP16 GS1-433O24.1 GS1-590J15.1 hsa-let-7f-2 hsa-mir-492 hsa-mir-98
HSD17B10 HUWE1 IQSEC2 LSM10 MED14 MID1IP1 MKRNP5 NAP1L2 NAV3 NYX
OSCP1 OTC PTCHD1 PTCHD1 RIBC1 RP11-126D17.1 RP11-126D17.4
RP11-157D23.1 RP11-157D23.2 RP11-169L17.2 RP11-169L17.3 RP11-169L17.5
RP11-185O17.3 RP11-204C16.4 RP11-265P11.1 RP11-265P11.2 RP11-272G22.1
RP11-272G22.2 RP11-272G22.3 RP11-320G24.1 RP11-40F8.2 RP11-469E19.1
RP11-478I12.1 RP11-494I9.1 RP11-494I9.2 RP11-540L11.2 RP11-654E17.2
RP1-169I5.4 RP11-77G22.2 RP11-77G22.3 RP13-126P21.1 RP13-13A3.1
RP13-444K19.1 RP3-339A18.3 RP3-339A18.6 RP4-646N3.1 RP4-646N3.3 RP5-
1172N10.2 RP5-972B16.2 RP5-972B16.2 RP6-186E3.1 RP6-29D12.2 RP6-29D12.3 RP6-
29D12.4 RPGR SLC6A15 SMC1A SNORA31 SNORA63 snoU13 snoU13 STK40 SYT1
THRAP3 TSPAN7 U4 U6 U6 U6 U7 USP9X Vault Y_RNA
Y_RNA ZXDB

PANTHER analysis: 30 mapped ids are found, 74 mapped ids are not found.

There is CnLOH present in ER and TN pure DCIS (n=017/17) not observed in HER2 positive pure DCIS (n=0/7).

1. p-values < 0.05;
2. A gene list is mapped from 84 genomic regions found on chromosomes 1, 2, 3, 4, 6, 7, 9, 10, 11, 12, 18, 20, 21;
3. These regions encompass 630 genes altered in ER pure DCIS and TN pure DCIS;

Genes:

7SK	AACS	ABCB11	ABHD11	ABTB2	AC007556.3	AC009336.1	AC009336.21	AC009336.23		
	AC009336.24		AC009475.2	AC010145.2	AC010145.3	AC010148.1	AC011998.2	AC011998.4	AC011998.5	AC016739.1
	AC016739.2	AC016757.1	AC017048.1	AC017048.2	AC017048.3	AC017048.4	AC026366.1	AC026366.2	AC055713.1	AC064874.1
	AC068706.1	AC068799.1	AC069209.1	AC069214.1	AC069234.1	AC073069.2	AC073846.1	AC073846.2	AC073846.3	AC078875.1
	AC090670.1	AC091609.1	AC091611.1	AC092162.1	AC092576.1	AC095030.1	AC096536.1	AC097713.2	AC097713.3	AC097713.4
	AC099680.1	AC104169.1	AC104387.1	AC104837.1	AC114486.1	AC114814.2	AC114814.3	AC114814.4	AC117503.2	AC122688.1
	AC125603.1	AC126309.2	AC127164.1	AC131263.2	AC132216.1	ACADS	ACOT11	ACTL8	ADAMTSL2	ADH5P2
	AFG3L2	AGAP1	AK3L1	AK5	AL031427.1	AL033381.1	AL033381.2	AL034344.1	AL035400.1	AL035706.1
	AL049710.1	AL049710.2	AL049710.4	AL049745.2	AL109843.1	AL136324.1	AL137855.1	AL138781.1	AL139244.1	AL157773.1
	AL158068.1	AL158167.1	AL160060.2	AL161452.1	AL161740.1	AL161915.1	AL162579.1	AL353771.1	AL353898.1	AL353898.2
	AL353898.3	AL353898.4	AL354946.1	AL355574.1	AL357054.1	AL358512.1	AL390994.1	AL391728.1	AL512329.1	AL512329.2
	AL512329.3	ALDH4A1	ANAPC5	ANKRD13C	ANXA10	AP000302.58		AP005264.2	ARHGEF10L	ARHGEF19
	ARL4C	ATP5GP1	ATP6V0A2	BAG2	BBS5	BCAS2	BEND6	BMP6	BRDT	BRI3BP
	BSND	BTBD8	C11orf74	C12orf43	C12orf49	C1orf134	C1orf141	C1orf168	C1orf175	C1orf177
	C1orf191	C1orf224	C1orf64	C1orf83	C1orf89	C2orf19	C6orf142	C6orf195	C8A	C8B
	C9orf116	C9orf62	C9orf69	CABP1	CAGE1	CAMKK2	CAMSAP1	CAPN10	CAPRIN1	CCDC73
	CCDC92	CDC7	CDCP2	CIDEA	CLCNKA	CLCNKB	CLDN3	CLDN4	COL5A1	COX6B1P7
	CR392000.1	CR392000.2	CRISP1	CRYZ	CST1	CST2	CST4	CST5	CSTP1	CSTP2
										CTH
	CYB5RL	DAB1	DBH	DCLRE1B	DDI2	DDX55	DDX60	DEFB110	DEFB112	DEFB113
	DEFB114	DEFB133	DENND2C	DEPDC1	DEPDC7	DHCR24	DHRS9	DHX37	DIO1	DIRAS3
	DNAH10	DNAJC30	DOK7	DSP	DST	EEF1E1	EIF2B1	EIF2S2P5	EIF3M	ELTD1
	ENTPD8	EPHX4	ESPNL	EVX2	EXD3	EXOC2	FAM101A	FAM131C	FAM132B	FAM151A
	FAM163B	FAM83B	FASTKD1	FBLIM1	FBXO21	FBXO3	FBXO42	FBXW8	FCN1	FCN2
	FOXC1	FOXF2	FOXQ1	FTLP17	G6PC2	GDI2P2	GFRAL	GLIS1	GLT6D1	GMDS
	GOT2L1	GPR177	GPR35	GREB1L	GTF2H3	HCRTR2	HECW2	HLA3	HIPK1	HNF1A
	HOXD1	HOXD10	HOXD11	HOXD12	HOXD13	HOXD3	HOXD4	HOXD8	HOXD9	HRK
										hsa-
mir-10b		hsa-mir-1262	hsa-mir-1290	hsa-mir-604	hsa-mir-938	HSP90B3P	HSPB11	HSPB7	HUS1B	IFFO2
										IFI44
	IFI44L	IFNGR2	IGSF21	IL12RB2	IL17A	IL23R	ILKAP	IRF4	KBTBD10	KCNT1
	KDM2B	KIAA0649	KIAA1586	KLHL30	KLHL31	KSR2	LASS6	LCN1	LCN9	LDLRAD1
	LHX3	LIN28AP3	LMO2	LPAR3	LPCAT2BP	LPHN2	LRP2	LRPAP1	LRRC1	LRRC40
	LRRC42	LRRC7	LYZL1	MACROD2	MAP1LC3B2	MCOLN2	MCOLN3	MIER1	MLEC	MRPL37

MRPS11P1	MRPS2	MTL1	MTX2	MUTED	MYCN	MYCNOS	MYLK4	NACC2	NAT10
NCOR2	NCRNA00118		NCRNA00159		NCRNA00173		NEGR1	NELF	NOS1
NOSTRIN	NOXA1	NRARP	OASL	OBP2A	OFCC1	OLFM1	OMA1	ORAI1	OSBPL6
P2RX4	P2RX7	PADI2	PAEP	PARS2	PAX7	PCSK9	PDE11A	PEBP1	PGAM3P
PGK2	PIGK	PIN1L	PIP5K1P1	PKHD1	PLEKHM2	PNPLA7	PPAP2B	PRIM2	PRKAA2
PRPF31	PRRG4	PTGFR	QSER1	QSOX2	RAB23	RAMP1	RAPGEF4	RBM45	RFC5
RILPL1	RIOK1	RNF34	RNFT2	RP11-101C11.1		RP11-109I2.3		RP11-124G5.1	
RP11-124I4.2		RP11-129K24.3		RP11-129K24.4		RP11-12C17.2		RP11-139O18.1	
RP11-145H9.3		RP11-157J24.1		RP11-159J16.1		RP11-169K16.4		RP11-169K16.6	
RP11-169K16.7		RP11-169K16.8		RP11-169K16.9		RP11-175G14.1		RP11-17E13.3	
RP11-180O5.2		RP11-181B18.1		RP11-192N10.2		RP11-203B9.4		RP11-203H2.1	
RP11-203H2.2		RP11-213P13.1		RP11-218C14.5		RP11-240D10.2		RP11-240D10.4	
RP11-240G22.1		RP11-243M12.1		RP11-263F14.3		RP11-288G3.2		RP11-288G3.3	
RP11-288G3.4		RP11-288I21.1		RP11-292O17.1		RP11-2B19.1		RP11-310I9.1	
RP11-320E2.1		RP11-324K6.1		RP11-335E14.1		RP11-338K17.1		RP11-339A11.1	
RP11-343L14.2		RP11-363H12.1		RP11-375A5.1		RP11-377K22.2		RP11-377K22.3	
RP11-378I13.1		RP11-380J14.1		RP11-380L11.1		RP11-393N21.1		RP11-397G17.1	
RP11-399H11.2		RP11-399H11.3		RP11-407P2.1		RP11-411K7.1		RP11-411K7.2	
RP11-411K7.4		RP11-411K7.5		RP11-426A6.5		RP11-426A6.7		RP11-426A6.8	
RP11-426A6.9		RP11-42O15.2		RP11-432J22.2		RP11-447M12.2		RP11-473A10.2	
RP11-475O6.1		RP11-478C1.6		RP11-478C1.7		RP11-478C1.8		RP11-47K11.2	
RP11-47K11.3		RP11-518D3.1		RP11-518D3.3		RP11-518D3.4		RP11-524H19.1	
RP11-524H19.2		RP11-524K22.1		RP11-529E10.8		RP11-532F6.2		RP11-534G20.3	
RP11-548C21.1		RP11-549I9.1		RP11-550H2.1		RP11-550H2.2		RP11-555H7.2	
RP11-5P18.1		RP11-5P18.10		RP11-5P18.2		RP11-5P18.3		RP11-5P18.5	
RP11-614L17.1		RP11-622P13.2		RP1-167A19.5		RP11-67L3.2RP11-67L3.4RP11-67L3.5RP11-67L3.6			
RP11-69L16.4		RP11-69L16.5		RP11-69L16.6		RP11-719J20.1		RP11-724O16.1	
RP11-779P15.1		RP11-779P15.2		RP11-79N23.1		RP1-181J22.1		RP11-82L20.1	
RP11-911I11.1		RP11-98L5.2RP11-98L5.4RP11-98L5.5RP1-20B11.2				RP1-28O17.1		RP1-303A1.1	
RP13-476E20.1		RP1-71H19.2		RP3-334F4.1RP3-334F4.2RP3-335N17.2		RP3-336K20__B.2		RP3-	
368B9.1	RP3-380B4.1		RP3-417L20.3		RP3-417L20.4		RP3-445O10.1		RP3-510B21.1
RP4-552O12.2		RP4-591B8.2		RP4-609E1.2		RP4-609E1.3		RP4-621B10.3	
RP4-621B10.3								RP4-	
633H17.2	RP4-641G12.3		RP4-641G12.4		RP4-654H19.2		RP4-660H19.1		RP4-677H15.2
RP4-677H15.4		RP4-694A7.2		RP4-694A7.3		RP4-694A7.4		RP4-705F19.1	
RP4-705F19.2		RP4-706A16.2		RP4-706A16.3		RP4-710M16.2		RP4-726F1.1RP4-733M16.2	
RP4-733M16.2								RP4-	
734C18.1	RP4-753D5.3		RP4-753D5.4		RP4-758J24.4		RP4-763G1.1		RP4-763G1.2
RP5-1033K19.2		RP5-1073O3.7		RP5-1077H22.1		RP5-1103B4.3		RP5-1177M21.1	
RP5-1177M21.1								RP5-	
827O9.1	RP5-831K15.1		RP5-836J3.1RP5-837I24.1		RP5-837I24.2		RP5-837I24.3		RP5-
RP5-837I24.4		RP5-837I24.5		RP5-837I24.6		RP5-866L20.1		RP5-866L20.2	
RP5-866L20.3									
RP5-997D24.3		RP5-997D24.5		RP6-102O10.1		RP6-239D12.1		RPE65	
RPS17P5		RPS5L	RSC1A1	SARDH	SCARB1	SCLY	SERBP1	SFRS11	RPS11P1
SLC35D1	SNORA31	SNORA58	SNORA70	SNORA9	SNORD112	SNORD115	snoU13	SNRNP48	SLC25A34
SLC25A34								SOHLH1	
SORBS2	SPC25	SPEN	SSBP3	SSR1	SSX2IP	ST6GALNAC3		ST6GALNAC5	
STX1A	SUDS3	SVIL	TAOK3	TAS1R2	TCEB1P18	TCP11L1	TCTN2	TESC	TFAP2B
TFAP2D	TGFR3	TINAG	TMED2	TMEM48	TMEM50B	TMEM59	TMEM61	TMEM82	TRIM33
TTC22	TTC4	TTLL7	TUBB6	TXNDC5	TYW3	U4atac	U5	U6	U6atac
U6atac								U7	
UBAC1	UBC	UBE2F	UNC119B	USP24	VPS37D	VSIG10	WBSCR22	WBSCR26	WDR63
WDR78	WSB2	XXyac-YX60D10.1		Y_RNA	YIPF1	ZBTB17	ZFYVE28	ZNF326	ZNF385B
ZNF451	ZNF664								

PANTHER analysis: 292 mapped ids are found, 338 mapped ids are not found;

5.9 Summary and Discussion of Genomic Differences between Pure DCIS and DCIS Associated with Invasive Breast Disease

This section addresses point 2 of the study aims (page 67): “To undertake an analysis of archival DCIS samples to identify genomic variation, abnormalities and changes between different pure DCIS lesions and DCIS lesions associated with invasion” and refers to pure ER positive DCIS compared to ER positive DCIS associated with invasive tumour (5.8.4) and pure TN DCIS compared to TN DCIS associated with invasive tumour (5.8.8)

Comparisons of copy number aberrations were used to produce charts highlighting molecular functions, biological processes and protein classes were generated. Examples are given within this chapter; however the majority of charts are saved to the accompanying disc.

In this initial analysis of the genomic data, only regions showing complete differences are examined. Caution should be taken in this series when looking at CnLOH in the samples analysed here due to the lack on corresponding normal tissues. Samples were normalised against a standard via Affymetrix, but not against associated normal tissues from the same sample.

5.9.1 *Genomic Differences found in Oestrogen Receptor Positive Pure DCIS Versus Oestrogen Receptor Positive DCIS Associated with Invasive Breast Disease*

(See section 5.8.4)

The frequency plots for ER pure DCIS and ER positive DCIS associated with invasion show a high degree of heterogeneity with ER pure DCIS having some homologous regions but mainly distinct and separate CNAs compared to ER positive DCIS associated with invasive breast disease. There are no amplifications present in ER positive pure DCIS whereas ER positive DCIS associated with invasive breast disease shows amplification found on chromosomes 6 and 8. CdLOH are varied, with ER pure DCIS showing loss of heterozygosity found on chromosomes 3, 5, 7, 9, 10, 11, 12, 11, 16, 17 and X. CdLOH is present on chromosome 5, 7, 10, 11, 12, 16, 17, 19, X in ER positive DCIS associated with tumour. ER positive pure DCIS shows duplication on chromosome 8 although this is not seen in the ER positive DCIS associated with invasive breast disease.

There are a large number of gains seen in ER positive DCIS associated with invasive breast disease across all chromosomes (except 13, X) whilst ER positive pure DCIS show gains less frequently on all, except certain regions found on chromosomes 1, 2, 4, 5, 6, 9, 11, 12, 16, 17, 18, 19 and 20. There are no gains on 12, 14 or 20.

As with pure TN DCIS versus TN DCIS associated with invasive breast disease, it may be that for ER positive lesions further genomic changes maybe either required or assist in the pathway to invasive disease.

There are amplifications identified in ER positive DCIS associated with invasive breast disease which we did not observe in ER positive pure DCIS. PANTHER analysis reveals 29 mapped genes. Of these, 13 are histones with epigenetic post translational modifications of histones being described as hallmarks of cancer (226). Elsheikh et al have correlated global histone modifications with breast tumour phenotypes, prognostic factors and patient outcome in invasive breast cancer. Utilising IHC on TMAs they identified; "Variations in bulk histone modifications in

different grades, morphologic types, and phenotype classes of invasive breast tumors. Furthermore, we have identified hypomodified and hypermodified tumor clusters, which correlate with known prognostic factors and clinical outcome” (227).

In this series there are duplications in the ER positive pure DCIS but not in ER positive DCIS associated with invasive breast disease. PANTHER analysis reveals a single gene, ASPH, which has been shown to be amplified in breast tumours in one previous study (228). Kadota et al (228) describe ASPH as one of four potentially novel oncogenes although they did not investigate ASPH further.

For gains in ER positive pure DCIS and ER positive DCIS associated with invasive breast disease there is considerable overlap in the gene patterns with no specific gains identified, and therefore these have not been further examined in this initial analysis.

Sc gains were found in ER positive pure DCIS not present in ER positive DCIS associated with invasive breast disease. PANTHER analysis reveals 8 mapped genes. Of these 2 are tyrosine kinases, one of these, tyrosine kinases (PTK2), has been previously shown to be up-regulated in DCIS (229). Also known as FAK Lightfoot et al (234) describe this kinase as a mediator of several functions including proliferation adhesion and survival. They found overexpression in DCIS in 66% of cases (n=34/51) and 33% of invasive lesion (n=6/18) but no expression in fibrocystic disease. They suggest that FAK has an active role in DCIS and survival in tumorigenesis. Other Sc gains were found in ER positive DCIS associated with invasive breast disease not observed in ER positive pure DCIS. PANTHER analysis reveals 62 known genes with a variety of molecular, biological and protein classes (see

PANTHER analysis) is used to create pie charts and radar map for molecular functions, biological processes and protein class (Figures 45-47).

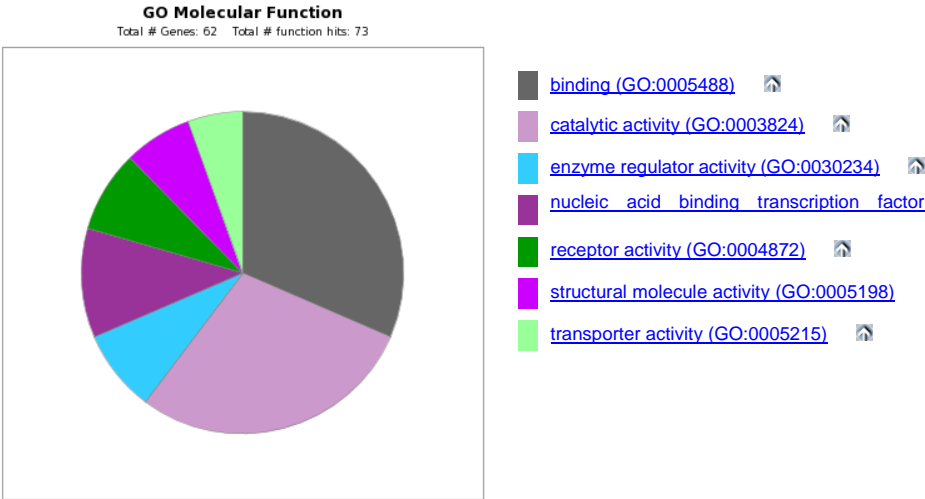


Figure 45: Molecular functions associated with Sc gains in ER positive DCIS associated with invasive breast diseases and ER positive pure DCIS.

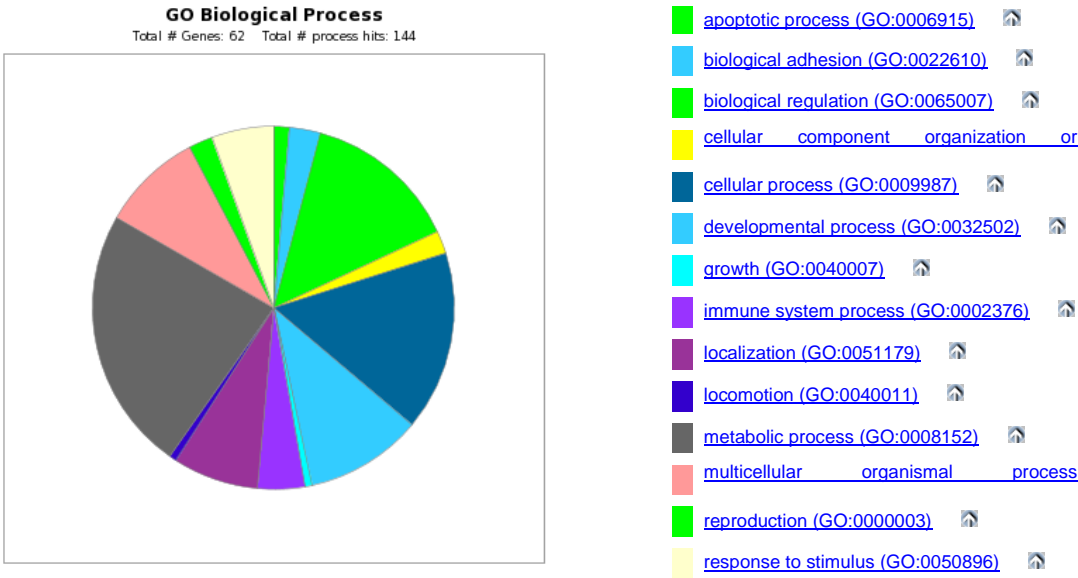


Figure 46: Biological processes associated with Sc gains in ER positive DCIS associated with invasive breast disease and ER positive pure DCIS

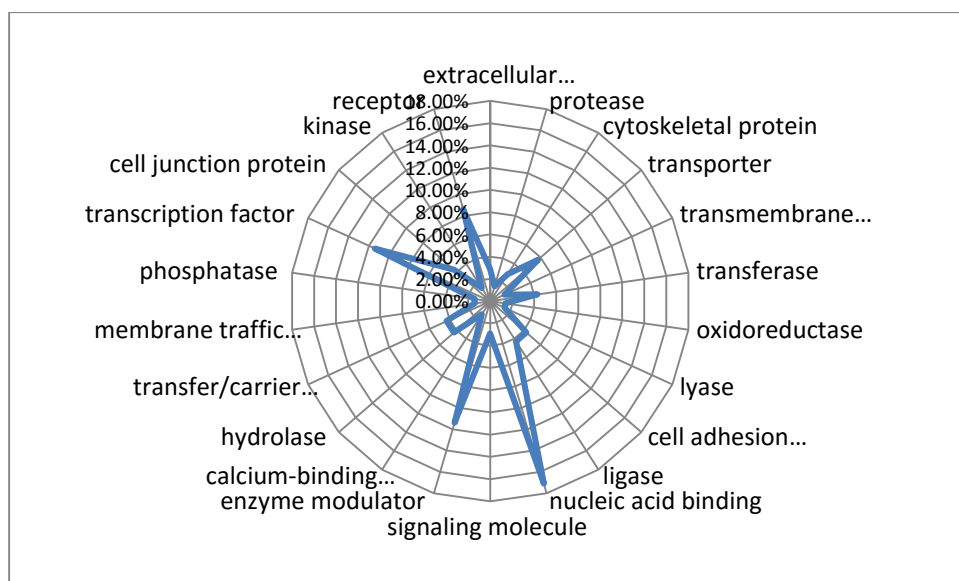


Figure 47: Protein classes associated with ER positive DCIS associated with invasive breast disease and ER positive pure DCIS.

Of the 62 genes, TNFRSF4 (a tumour necrosis factor) has been previously described in both invasive disease and DCIS (230). Also known by the alias OX40, this membrane-bound member of the tumour-necrosis-factor-receptor (TNFR) superfamily, plays an important role in proliferation, survival and infiltration of activated T cells via binding to OX40L (230). Xie et al suggest that high OX40 expression may be associated with malignant transformation, progression, invasion and metastasis in breast cancer biology (230).

There are losses in ER positive DCIS associated with invasive breast disease not observed in ER positive pure DCIS. PANTHER analysis reveals 42 mapped genes, of which 6 are zinc finger proteins. Losses include NRG3 which encodes for the EGF containing ligands that mediate binding to ERBB receptor tyrosine kinases (231). NRG is reported to regulate mammary phenotype (232).

Total losses in pure ER positive DCIS include RB1 (retinoblastoma gene) that has tumour suppressor function. It has been shown to be frequently altered in breast carcinomas leading to loss of expression (233). Total Losses seen in DCIS associated with invasive breast disease but not in pure DCIS are far more common (Figures 49-51). PANTHER analysis reveals 219 mapped genes. There is a wide variation in the genes and the molecular functions with total loss. Of these MAP2K4 (mitogen activated protein kinase) whose deletion has been suggested as one of a set of putative cancer genes (234).

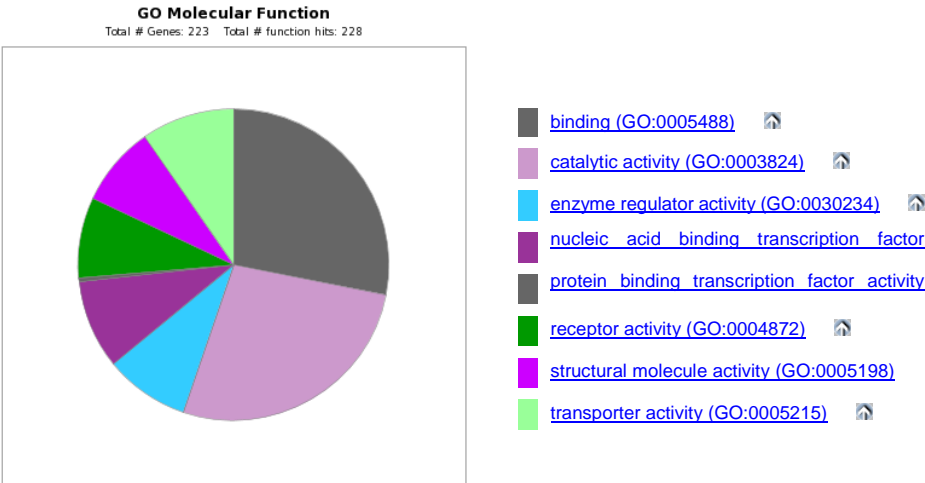


Figure 48: Molecular functions associated with total loss in ER positive DCIS associated with invasive breast disease.

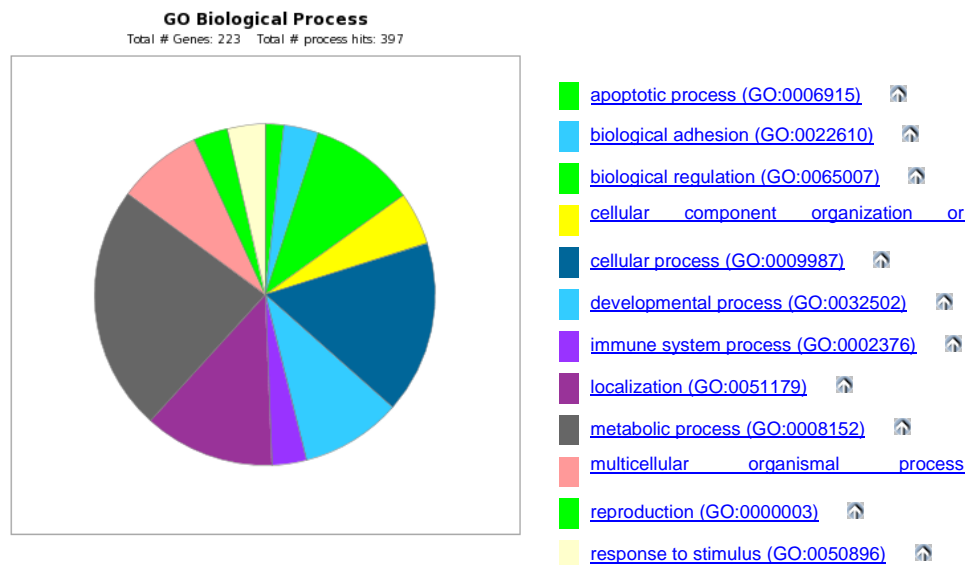


Figure 49: Biological processes associated with total loss in ER positive DCIS associated with invasive breast disease

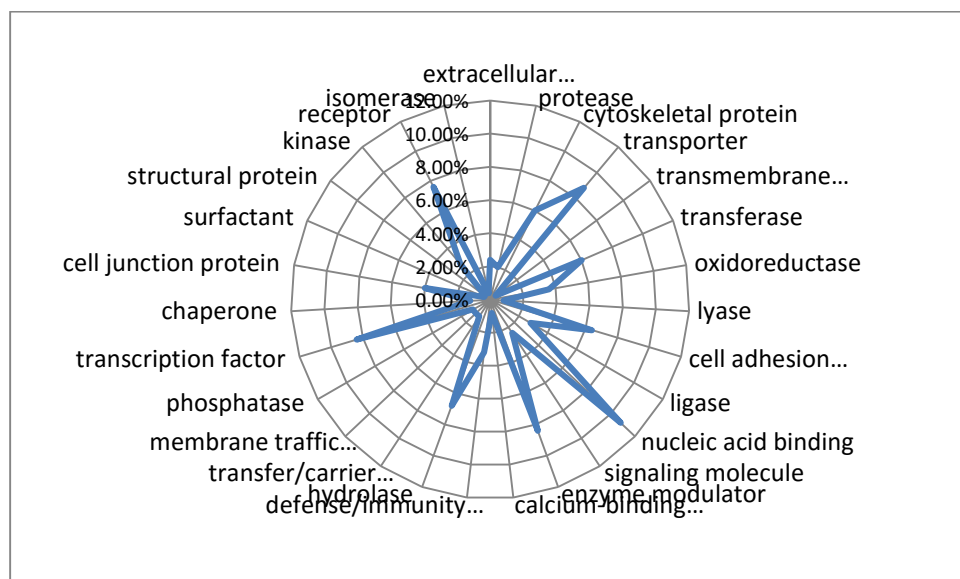


Figure 50: Protein classes associated with ER positive DCIS associated with invasive breast disease.

There is CdLOH present in ER positive pure DCIS not observed in ER positive DCIS associated with invasive breast disease. PANTHER analysis reveals 176 mapped

genes with a variety of molecular functions and protein classes. A number matrix metalloproteases or MMPs (MMP1, MMP3, MMP 7, MMP8, MMP10, MMP13, MMP20, MMP27, PTGFD) have CdLOH ER positive pure DCIS. Loss of Matrix metalloproteases (MMPs) activity may result in a wide range of diseases including cancer (235).

CdLOH present in ER positive DCIS associated with invasive breast disease has similarities to the ER pure DCIS with no specific changes identified and is not examined further in this initial analysis.

There is CnLOH present in ER positive pure DCIS not observed in ER positive DCIS associated with invasive breast disease. PANTHER analysis reveals 248 mapped genes

representing a wide variety of molecular functions (Figures 51-53). ERBB 2, MDM2 (a regulator of p53) and Notch2 have CnLOH. MDM2 (236) and Notch2 (237) have been suggested as candidate genes in breast cancer progression.

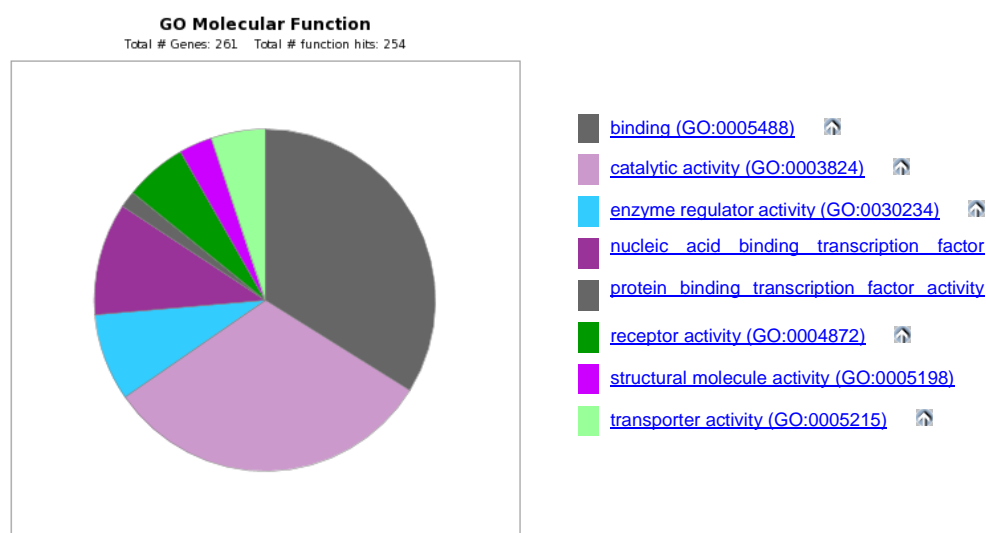


Figure 51: Molecular functions associated with CnLOH in ER positive pure DCIS

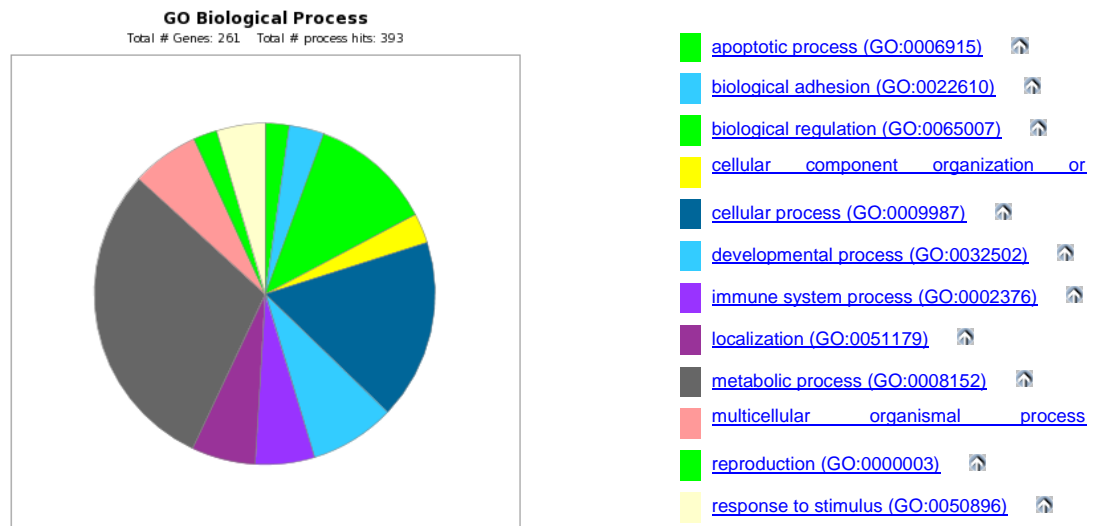


Figure 52: Biological processes associated with CnLOH in ER positive pure DCIS

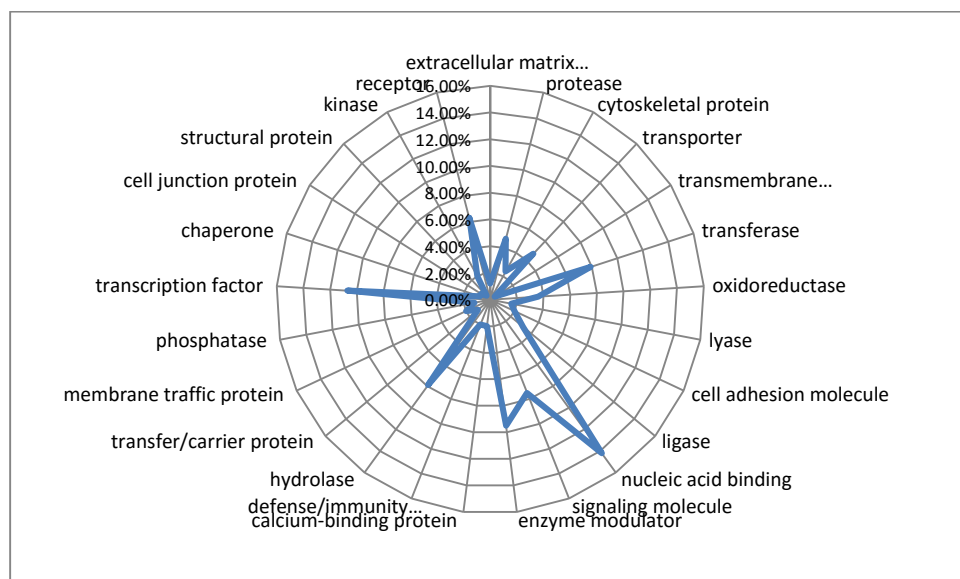


Figure 53: Protein classes associated with CnLOH in ER positive pure DCIS

There is CnLOH in ER positive DCIS associated with invasive breast disease which is not observed in ER positive pure DCIS. PANTHER analysis reveals 64 mapped

genes. In this series the gene S100z shows CnLOH whereas in one previous study it was reported to be up-regulated (238).

5.9.2 Genomic Differences found in Triple Negative Pure DCIS Versus Triple Negative DCIS Associated with Invasive Breast Disease

The frequency plots for CNAs show that samples of triple negative DCIS associated with invasion in this series exhibit many more genomic aberrations at the chromosomal level than pure triple negative DCIS. CdLOH, gains, losses and Sc gains all show a higher frequency of aberrations for the DCIS associated with invasive breast disease than pure DCIS which has not progressed. There are some homology present with CnLOH (see below) and total losses in TN pure DCIS and TN DCIS associated with breast disease. CdLOH are also located on distinctly separate chromosomal and loci for pure TN DCIS (chromosome 7, 16) compared to CdLOH on chromosome 3, 4, 5, 8, 9, 13, 14, 15, X for TN DCIS associated with invasive disease. Aberrations in both are located on chromosome 17 but in different regions. Gains are present found on chromosomes 1 and 8 for both types of lesion, but at a lower frequency for TN pure DCIS. Gains are found only on chromosome 3, 6, 11, 13, 18, 21 for TN DCIS associated with invasive breast disease. Losses show a similar pattern, with pure TN DCIS having distinct losses on chromosome 7, 16, and 19 compared to distinct losses on 4, 5, 8, 9, 14, X in TN DCIS associated with invasive breast disease. Sc gains on chromosome 1, 2, 5, 6, 8, 9 are more abundant in TN DCIS associated with invasive breast disease with pure DCIS only showing Sc gains on chromosome 1.

These differences in TN pure DCIS and TN DCIS associated with invasive breast disease may indicate that although a cancerous phenotype exists within pure TN DCIS several further copy number aberrations may be necessary to switch to an

invasive TN phenotype. Both gains and losses are present in TN DCIS associated with invasion which are not seen in pure TN DCIS in this series. There is a degree of homology in some chromosomal changes such as gains on chromosome 1 and losses on chromosome 4 in both pure TN DCIS and TN DCIS associated with invasive breast disease. However, there are few gains and losses present found in pure DCIS alone.

Comparison of TN pure DCIS and TN DCIS associated with invasive breast disease shows no amplifications or duplications found in either process.

Comparing the chromosomal gains in TN pure DCIS and in TN DCIS associated with invasion there is some homology present, with pure TN DCIS cases having gains comparable to those seen in the TN DCIS associated with invasion in 5/9 cases. However, the TN DCIS associated with invasion shows additional gains not present in the TN pure DCIS. PANTHER analysis of the genes identifies 16 known genes with a range of biological functions associated with several protein classes (Figure 54-56). Of these, one - NFIB has been previously identified as associated with ER negative invasive breast cancer (239) and proposed as a genomically distinct TN subgroup mainly composed of adenoid cystic carcinomas (240).

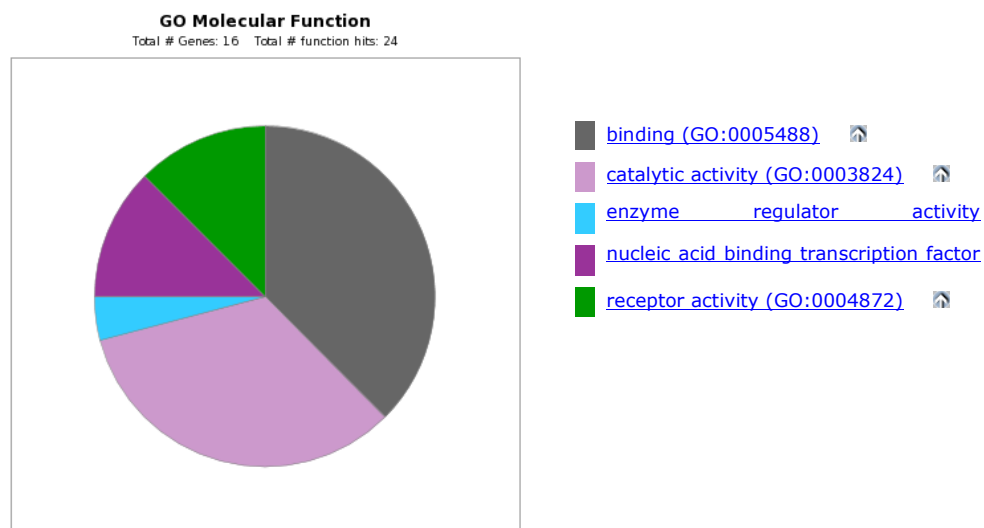


Figure 54: Molecular functions associated with the genomic gains identified in TN DCIS associated with invasive breast cancer.

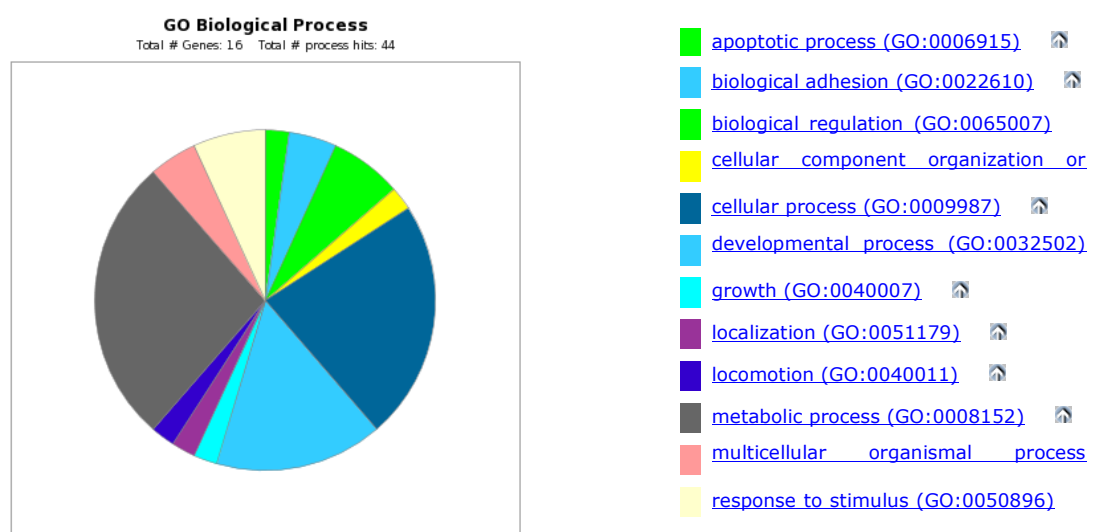


Figure 55: Biological processes associated with the gains identified in TN DCIS associated with invasive breast cancer.

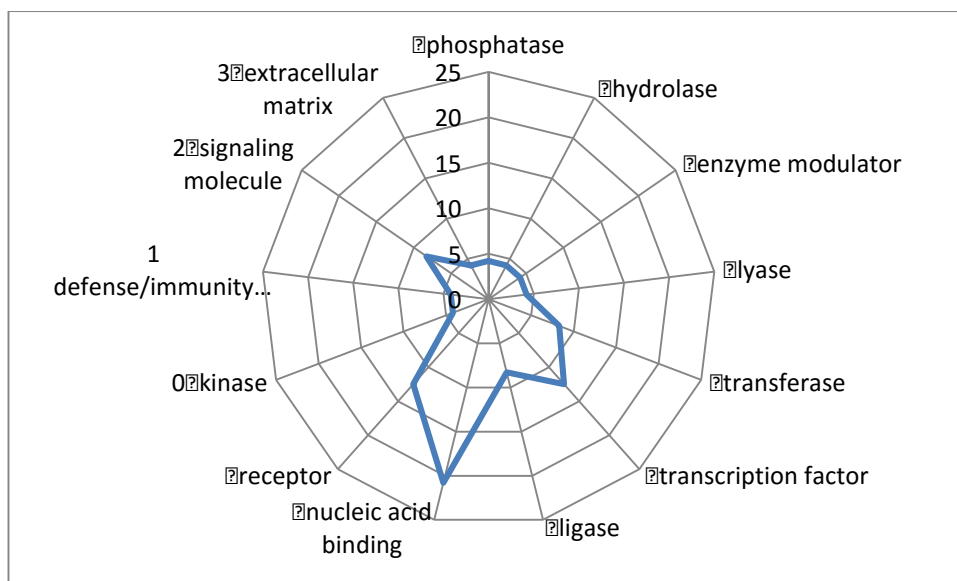


Figure 56: Protein classes associated with the genomic gains identified in TN DCIS associated with invasive breast cancer.

Sc gains were present in TN DCIS associated with invasive breast disease but not in TN pure DCIS. PANTHER analysis maps 86 of the 350 genes giving a range of molecular and biological functions associated with several protein classes (Figure 57-59)

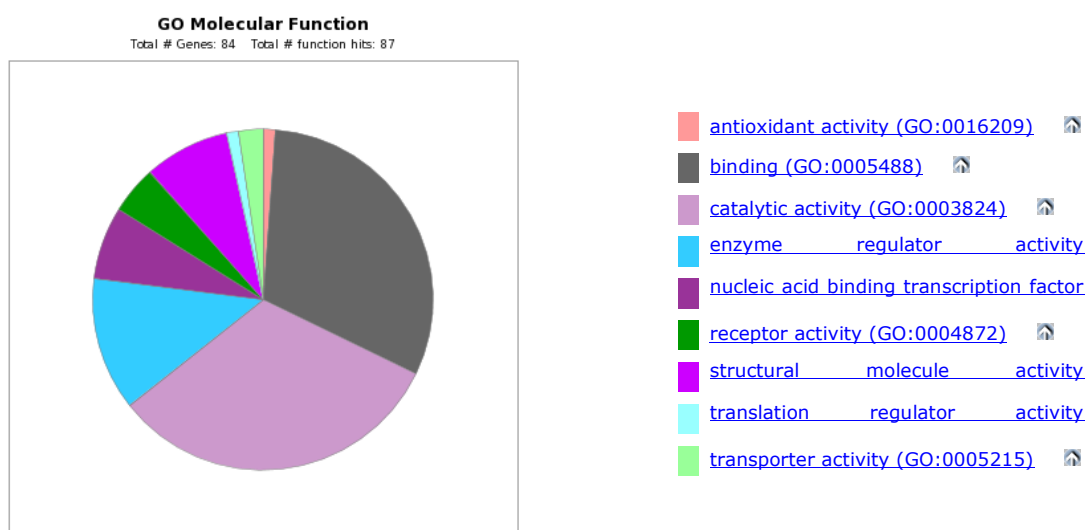


Figure 57: Molecular functions associated with Sc gains for TN DCIS associated with invasive breast disease.

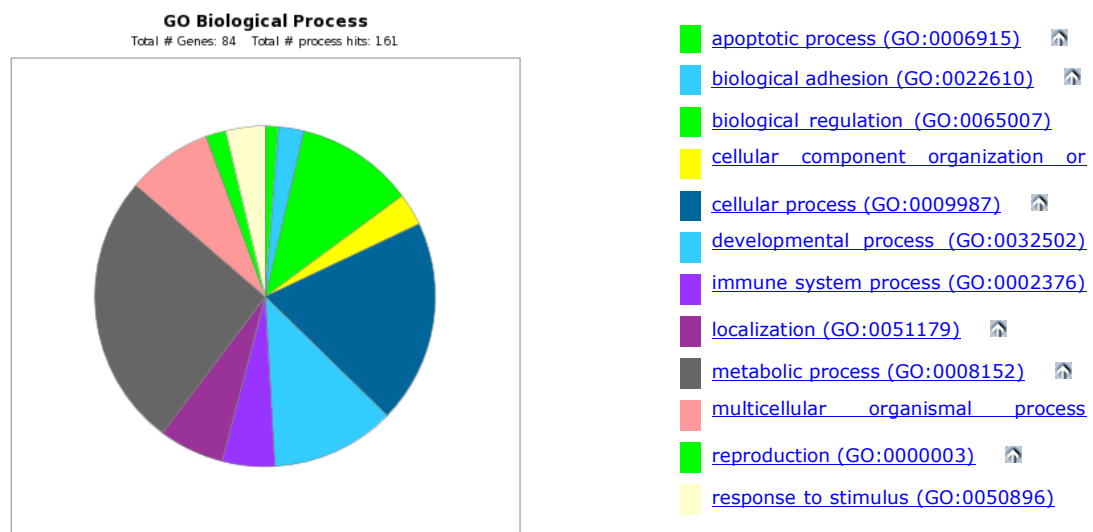


Figure 58: Biological processes associated with Sc gains for TN DCIS associated with invasive breast disease.

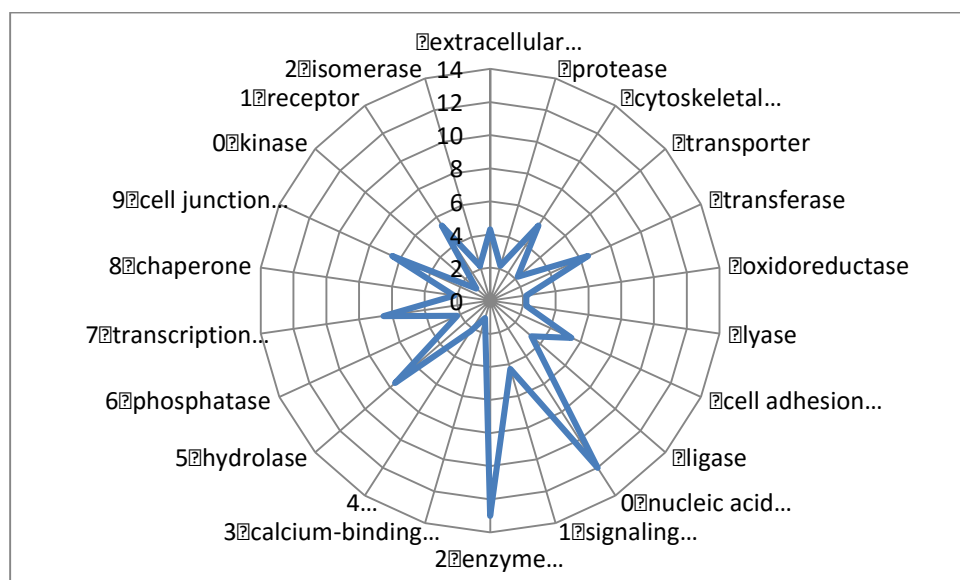


Figure 59: Protein classes associated with Sc gains in TN DCIS associated with invasive breast disease.

The genes showing Sc gains in TN DCIS associated with invasive breast disease but not observed in TN pure DCIS include PTP1B (Protein tyrosine phosphatase 1B) which has been reported to have both tumour suppressor and tumour promoting roles (241). Similarly ADAMTS-12, a secreted metalloprotease, shows both oncogenic and

tumour-suppressive effects has Sc gains in this series (242). PLK3 also shows Sc gain; this is a member of the pololike kinases and whilst is involved in cell cycle regulation its expression remains steady throughout the cell cycle. It has been shown to be widely expressed in cancer cell lines (243). Some cadherin's (CDNH) 6 (K-cadherin), 9 (T1 cadherin) 12 (n-cadherin), 18 (ungrouped) show Sc gains in this series.

As opposed to the Sc gains seen in the TN DCIS associated with invasive compared to the pure form of TN DCIS, losses in TN pure DCIS not observed in TN DCIS associated with invasive disease are noted, with six genes mapped by PANTHER analysis. Three of these are in the type 1 keratins (14, 16, 17) which are structural components of the cytoskeleton.

Conversely, 7 gene losses are mapped by PANTHER analysis in TN DCIS associated with invasive breast disease showing a variety of molecular functions.

Total losses for pure DCIS not observed in TN DCIS associated with invasive disease reveal 40 genes mapped by PANTHER analysis. These include BRCA1, zinc finger like proteins (333 and 558) and olfactory receptor (OR) gene. Previous studies have shown up-regulation of ten OR genes in breast cancer cell lines clustered in 11q12.1 (244).

Total losses for TN DCIS associated with invasive breast but not observed in TN pure DCIS associated with invasive disease reveal 46 genes mapped by PANTHER analysis. These include metalloproteases, protease zinc finger protein, and extra cellular matrix protein. In this series there is total loss of the apoptotic marker PDCD6IP (3p23) in some cases of TN DCIS. Previously copy number analysis has reported that amplification of this gene is associated with an increased risk of

recurrence in invasive breast disease when considered as part of a trio of markers (*CYP24/PDCD6IP/BIRC*) for ER/PR-positive cancers. Individually *PDCD6IP* is reported to have no significant correlation in invasive breast cancer prognosis but does when associated with *CYP24* and *BIRC* (245) highlighting the complex interactions of different genomic markers.

CdLOH for TN DCIS associated with invasive disease not observed in TN pure DCIS revealed 12 genes mapped by PANTHER analysis. These include a kinase suppressor of Ras1 (*KSR1*) which has been identified as a potential tumour suppressor in *BRCA1* tumours (246). CnLOH for TN DCIS associated with invasive disease show similarities in 5/9 cases with the CnLOH in pure DCIS. However, there are 10/10 cases showing CdLOH observed in TN pure DCIS not observed in TN DCIS associated with invasive disease. PANTHER analysis reveals genes mapped by PANTHER analysis. Conversely, CdLOH for TN pure DCIS not observed in TN DCIS associated with invasive disease revealed 4 altered genes mapped by PANTHER analysis of which 3 are keratins -14, 16,17.

5.7.3 Observations on Genomic Analysis of Pure DCIS Compared to DCIS Associated with Invasive Breast Disease.

In this series the following observations can be made regarding analyses:

Amplifications and duplications between pure DCIS and DCIS associated with invasive disease show no similarity regardless of subtype.

Gains and losses show similarity for all types of DCIS with regard to chromosomal region, however the number of genomic gains or losses in DCIS associated with tumours is greater than those found in pure DCIS. In addition there are gains present in tumour associated DCIS not seen in pure DCIS lesions. This would fit with a hypothesis that further downstream genetic mutations occur post DCIS and lead certain credence to the possibility that not all pure DCIS is not wholly committed to an

invasive phenotype. Downstream events post establishment of pure DCIS may be required to for the initiation of invasive potential. SC gains show no overall pattern between these two cohorts.

Total losses between pure DCIS and DCIS associated with tumour show few matching genetic aberrations. As with amplifications and duplications and Sc gains these differences may indicate drivers in pure DCIS that inhibit the progression to invasive disease or in DCIS associated with invasion they may indicate invasive promoting genes. An alternative approach is that these differences could indicate increased heterogeneity giving rise to unidentified subtypes.

Copy neutral and copy deletion loss of heterozygosity is marked in both sets of data regardless of tumour subtype. The use of MIP arrays used control tissue of normal human breast as a reference point to give a baseline for the samples analysed. Ideally the control reference would have been taken from normal tissue from the same patient. The lack of normal tissue availability prevented this approach so it is acknowledged that there may be a certain amount of background “noise” which gives a higher number of comparative genomic differences than expected as there would be a disparity between the normal tissue of different individuals.

5.10 Summary of Genomic Differences between DCIS and Invasive Breast Disease

This section addresses point 4 in the study aims: “To identify any protein and/or genomic changes in all DCIS (pure and that associated with invasion) compared to invasive breast disease to determine potential biomarkers responsible for progression to an invasive state”.

In this series only those copy number aberrations seen in one series and absent in the second set, i.e. present in pure DCIS but not in invasive carcinoma, and vice versa, are addressed in this initial analysis.

5.10.1 Genomic Difference between ER Positive DCIS and ER Positive Invasive Breast Disease. (See section 5.8.4).

The frequency plots for CNAs in all ER positive DCIS compared to ER positive invasive disease show a mixed pattern, with some homology and a degree of heterogeneity. CdLOH shows similar aberrations on regions of chromosomes 8, 10, 11, 13, 15, 19, 20 X for both sample sets although ER positive DCIS has a lower frequencies. However, there are exclusive CnLOH regions found on chromosomes 1, 3, 4, 7, 10, 14, 15, 18, X in ER positive invasive breast cancer not seen in ER positive DCIS. There are duplications seen only in ER positive invasive disease, found on chromosomes 1, 8, 9, 10, 17, 21 not present in ER positive DCIS. The gains found show homology on chromosome 6 and 16 with a lower frequency in DCIS. There are losses on chromosomes 1, 2, 5, 7, 8, 11, 13, 15, 16, 19, 21 X found in both sample sets with a lower frequency found in ER positive DCIS. The exception is chromosome 11 where ER positive DCIS samples have greater losses than ER positive tumours. Sc gains are seen at similar frequencies but, as with other CNAs, the frequency in DCIS is lower. For total loss there are many more aberrations in ER positive invasive breast disease than ER positive DCIS. Again, there are some homologous regions (chromosomes 2, 6, 8, 8, 10, 13, 14, 16, 17, x) and regions of exclusive total loss for ER breast disease (chromosomes 2, 3, 4, 5, 7, 8, 9, 10, 12, 13, 14, 15, 16, 19, 20, 21, X). These data all indicate that further genomic differences are present in invasive breast disease not present in DCIS.

There are no amplifications in the comparison of ER positive DCIS versus ER positive invasive breast disease.

Duplications found in ER DCIS are also present in ER positive invasive breast disease and are not addressed in this initial analysis. Conversely, there are duplications in ER positive invasive disease not present in ER positive DCIS. PANTHER analysis reveals 137 mapped genes. In this series the gene loss or CnLOH of PTPRD is found in both pure ER positive DCIS and ER DCIS associated with invasion. In this series Muc1 is also duplicated; this has previously been shown to be elevated in ER positive invasive breast disease (247) but having less frequent expression (i.e. fewer cases) in ER positive DCIS (248). Notch2 also show duplications (236).

There are genomic similarities in the pattern of gains, Sc gains and losses in ER positive DCIS and ER positive invasive breast disease which have not been analysed in this preliminary study.

There are total losses of genes present in ER positive DCIS that are not present in ER positive invasive breast disease. PANTHER analysis reveals 40 genes showing total loss. Total loss in this series includes BRCA1 and several olfactory receptor genes.

CdLOH found in ER DCIS are also present in duplications for ER positive invasive breast disease and are not examined in this thesis.

There is CnLOH present in ER positive DCIS not observed in ER positive invasive breast disease. PANTHER analysis reveals 42 mapped genes. Of these genes MACROD2 has been suggested as a possible cause of Tamoxifen resistance when

overexpressed in metastatic tumours (249). PANTHER analysis of the converse, CnLOH present in ER positive invasive breast disease not observed in ER positive DCIS, reveals 37 mapped genes. Of these genes TIMELESS, one of the core circadian genes, has been proposed as an epigenetic risk factor by possibly regulating hormone functions. Its overexpression is reported to influence breast carcinogenesis (250) and warrants

5.10.2 Genomic Difference between Triple Negative DCIS and Triple Negative Invasive Breast Disease.

The frequency plots for CNAs show that there is a high degree of homology between TN DCIS and TN invasive tumours. DCIS lesions show similar but less complex copy number aberrations to the invasive disease. This concurs with the findings by Johnson et al who performed a similar MIP analysis on 21 cases of invasive ductal carcinoma with synchronous DCIS although with no phenotypic subgrouping (251). In the present series, a high degree of total loss is apparent found on chromosomes 3, 13, 14, 15 and 17. Gains found on chromosomes 3 and 6 are also identified in both TN DCIS and in TN invasive carcinoma, whilst gains in TN invasive breast cancer are seen in chromosome 11 although this is not apparent in TN DCIS. CdLOH is seen in both TN DCIS and TN invasive breast disease found on chromosomes 4, 12, 18 and X. However, for CnLOH TN DCIS shows a higher degree of complexity than TN invasive disease. CnLOH for TN DCIS are present more frequently found on chromosomes 1, 2, 3, 6, 7, 8, 9, 10, 11, 12, 19, 20, and X. CNLOH is present in TN DCIS found on chromosomes 5, 13, 16 and 21, where no CnLOH is found in TN invasive breast disease.

Gains and (and Sc gains) show very similar patterns between DCIS and invasive breast disease with regard to chromosome number. The level of complexity is raised in the invasive samples.

There are no amplifications in the comparison of TN DCIS versus TN invasive breast disease.

There are duplications found in TN invasive breast cancer not present in TN DCIS. PANTHER analysis reveals 62 mapped genes. Of these, protein tyrosine kinases PTPRK and PTK7 show duplications. PTPRK has been reported as a potential tumour suppressor (252) whilst positive immunohistochemical expression of antibodies raised against PTK7 has been shown in anthracycline-resistant invasive breast disease (253).

Gains, Sc gains and losses in TN DCIS and TN invasive disease show overlap of genes and are not examined in this initial analysis.

There are total gene losses found in some cases of TN DCIS not observed in some cases of TN invasive disease. PANTHER analysis reveals 205 mapped genes with a variety of molecular functions, biological processes and protein classes (Figures 60-62).

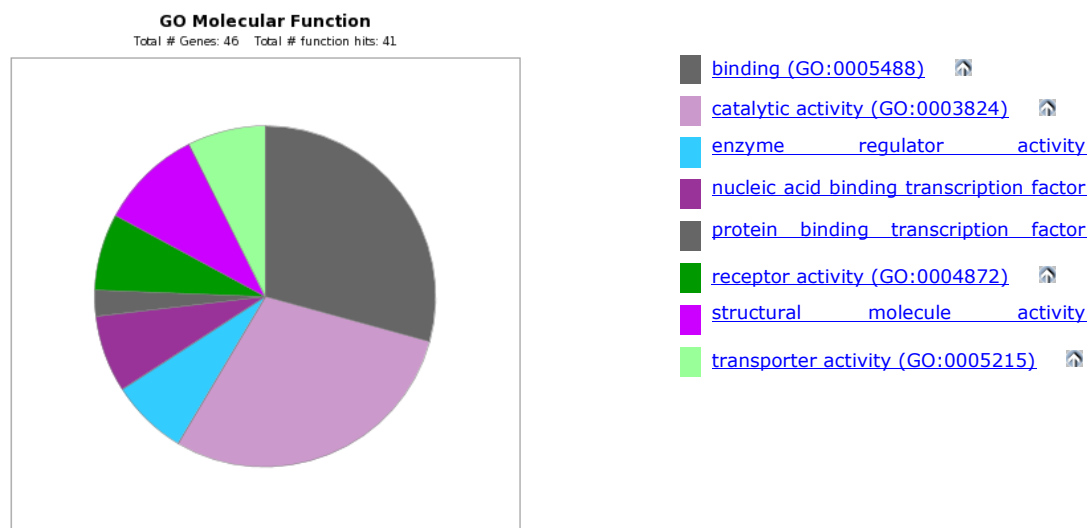


Figure 60: Molecular functions associated with total loss in TN DCIS

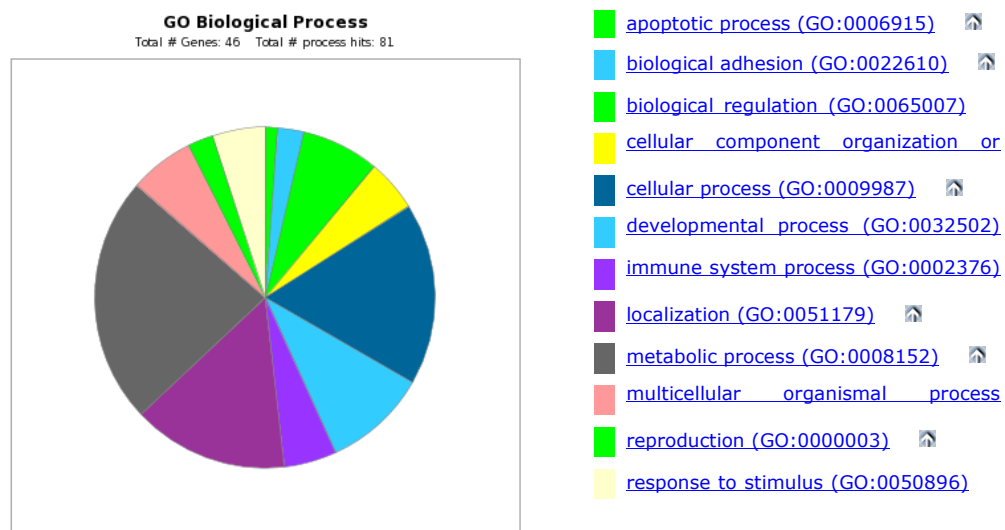


Figure 61: Biological processes associated with total loss in TN DCIS

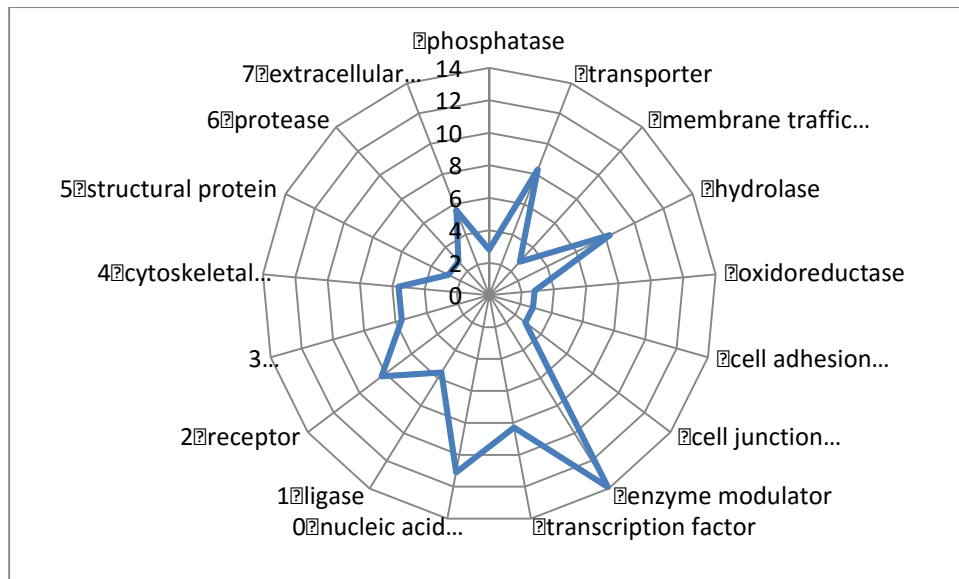


Figure 62: Protein classes associated with total loss in TN DCIS

In this series, FOXD1 shows total genomic loss in some cases of TN DCIS whilst it has been suggested as an upregulated oncogene in breast cancer (254).

There are total losses found in some cases of TN invasive breast disease not seen in some cases of TN pure DCIS. PANTHER analysis reveals 160 mapped genes. Olfactory receptor genes are predominant with 11 showing total loss (OR1B1, OR1J2, OR1J4, OR1L1, OR1L3, OR1L4, OR1L6, OR1L8, OR1N1, OR1N2, OR1Q1). Fonseca-Sanchez et al identify 55 olfactory receptor genes up-regulated in breast cancer cell lines (244). However, the olfactory receptor gene family is the largest in the genome.

There is CnLOH present in TN DCIS that are not observed in TN invasive breast disease. PANTHER analysis reveals 205 mapped genes. Genes showing high levels of CNLOH in invasive breast cancer include, CSNK2B and HLA-DQA1. CSNK2B is a casein kinase a ubiquitous protein kinase which regulates metabolic pathways, signal transduction, transcription, translation, and replication. Kren et al (255) suggest that

down-regulation of casein 2 induces loss of breast cancer cell viability and may be a potential target for triple negative breast cancer therapy. HLA-DQA1 (major histocompatibility complex, class II, DQ alpha 1) is one of the HLA (human leucocyte antigen complex). It encodes for a protein of the major histocompatibility complex (MHC II) which is involved in antigen presentation within the immune system. A variant of HLA-DQA1 has been implicated in increased risk of breast cancer (256).

There is CnLOH present in TN invasive breast disease not observed in TN DCIS. PANTHER analysis reveals 140 mapped genes. Of these, there is CnLOH of a number of taste receptor genes (TAS2R10, TAS2R13, TAS2R14, TAS2R19, TAS2R20, TAS2R31, TAS2R42, TAS2R50, TAS2R7, TAS2R8, TAS2R9). There is no current literature on the significance of these in breast cancer.

5.10.3 Observations on Genomic Analysis of DCIS Compared to DCIS Invasive Breast Disease.

In this series the following observations can be made which are true for all DCIS compared to invasive tumour regardless of subtype (see 5.8.6. and 5.6.10) No amplifications are present for DCIS compared to invasive breast disease. Duplications are more prevalent in invasive breast disease indicating more genomic instability. Invasive breast disease exhibits a greater number of gains and SC gains compared to DCIS further illustrating more complex genetic aberrations. There are also greater losses in invasive breast disease. There is a similar pattern observed in the chromosomal regions of duplications, gains, losses which could in turn be explained as a sampling bias. DCIS and invasive samples from the same patient were included in this cohort and not matched against each other in this analysis. This would result in similar chromosomal regions of instability if the DCIS had similar genetic

aberrations to the associated invasive tumour analysed. The discrepancy in lower signal could be due in part to the pure DCIS samples not showing the same genomic changes as DCIS associated with invasive breast disease. Interestingly cnLOH (unlike cdLOH) shows greater genomic variation in DCIS than invasive breast disease. It may be that these multiple minor variations reduce as clonal expansion takes place of tumour cells with a proliferative advantage. Further research on this dataset is required to illicit the key drivers of these results.

5.11 Summary of Genomic Differences between Pure DCIS versus Invasive Breast Disease

5.11.1 Pure ER Positive DCIS compared to ER Positive Invasive Breast Disease

The frequency plots for CNAs between ER positive pure DCIS and ER positive invasive breast disease exhibit some homologous regions for both gains and losses when compared to invasive breast disease, albeit with much higher frequencies seen in ER positive invasive breast cancers. Gains on chromosome 6, 8 16, 20 and 21 are mirrored between both sets. Similarly losses on regions of chromosomes 8, 9, 11 12, 15, 16, 19, X are also present in both types of sample in this series. There are a greater number of losses on several regions of chromosome 11 in ER positive pure DCIS. Sc gains in this series show little homology, with Sc gains found on chromosomes 1 and 8 in ER pure DCIS and Sc gains found on chromosomes 6 and 20 in ER positive invasive breast disease. There are Sc gains on chromosome 8 for ER pure DCIS also seen in ER positive invasive breast disease. There are many more CNAs showing total loss in ER positive invasive breast disease than ER positive pure DCIS.

These finding somewhat contradict the findings seen in the comparison of ER positive pure DCIS and ER positive DCIS associated with invasive breast cancer. ER positive invasive breast cancer samples compared to ER positive pure DCIS seem to have

more in common than pure DCIS and DCIS associated with invasive breast disease. It must be acknowledged that this is a small cohort (ER positive pure DCIS n=8, ER positive invasive breast disease n=9) and this may contribute. However, it may be that combinations of genomic changes are required in the progression from DCIS to an invasive phenotype. There may be genomic changes already present in pure DCIS that are required for tumour survival that are uninvolved in progression to an invasive state. These may be conserved in an invasive tumour but play no role in invasion. Conversely these genomic changes may be required to trigger downstream mutations in order for invasion to occur and be conserved as either drivers of invasion or redundant post DCIS mutations.

There are no amplifications in ER positive pure DCIS in this series. There are amplifications in ER positive invasive breast disease. PANTHER analysis reveals 1 gene (*EIF3H*) which has previously been significantly associated with risk of low-grade breast cancer (257). Located on chromosome 8q23 the exact role of E1F3H in cell cycle regulation is unknown(258). The EIF3H gene encodes the eukaryotic translation initiation factor 3 (EIF-3) complex, associated in several steps in protein synthesis initiation and mRNA recruitment. Translational control is a crucial component of cancer development and progression (259), and EIF3H in particular is frequently amplified in breast and prostate cancers (260).

No duplications are found in either ER positive pure DCIS or ER positive invasive breast disease.

There were gains present in both ER positive pure DCIS and ER positive invasive breast cancer and, without significant differences, these are not addressed in this thesis. However, there were gains found in ER positive invasive breast disease not found in ER positive pure DCIS. PANTHER analysis reveals 26 mapped genes. Of

these 26 genes, SPIRE1 has been proposed as a possible driver of invasion as under- or over-expression can result in increased or decreased extra cellular matrix degradation (261).

There were Sc gains present in both ER positive pure DCIS and ER positive invasive breast disease which are not analysed in this preliminary assessment.

There are losses present in ER positive pure DCIS and ER positive invasive breast disease not addressed in this initial analysis.

There are total losses of genes which are present in ER positive pure DCIS but not seen in ER positive invasive breast disease. PANTHER analysis reveals 13 mapped genes. Of these, NOX4 is proposed as a contributory gene in the epithelial to mesenchymal transition (EMT) process involved in cancer invasion (262).

There are genes showing total loss in ER positive invasive breast disease not found in ER positive pure DCIS. PANTHER analysis reveals 172 mapped genes. GBJ2 is a gene responsible for cellular adhesion and has also been shown to be present in invasive disease and proposed as a possible cause for invasion, although this is contradicted by the present findings (263).

There is CdLOH present in both ER positive pure DCIS and ER positive invasive breast disease not analysed in this thesis. Similarly there are CnLOH present in both ER positive pure DCIS and ER positive invasive breast disease which are omitted from this preliminary work.

5.11.2 Pure Triple Negative DCIS versus Triple Negative Invasive Breast Disease (See section 5.8.9).

The frequency plots for pure TN DCIS compared to TN invasive breast disease exhibit many more CNAs in the invasive breast disease compared to DCIS although some homology is present the majority of changes exist in TN invasive breast disease

alone. The genomic frequencies also show a higher complexity of genomic aberrations across multiple chromosome regions compared to corresponding ER positive DCIS and invasive breast disease. This reveals a much more complex phenotype, which may possibly contain sub-groups within this series. CdLOH is found exclusively on regions of chromosomes 4, 5, 8, 9, 10, 12, 13, 14, 15, 16, 18, X for TN invasive breast cancer. Shared regions, although at much lower frequency, are found on chromosomes 4, 5, 14 and 15. Exclusive regions of gain are seen on chromosomes 1, 3, 6, 8, 9, 10, 13 and 20. Sc gains are seen exclusively in TN invasive breast carcinomas in this series found on chromosomes 1, 3 and 8.

These findings are similar to those seen in TN pure DCIS versus TN DCIS associated with invasive breast disease indicating that genomic changes found in invasive breast disease are much more complex than those seen in DCIS.

There are no amplifications in TN pure DCIS or TN invasive breast cancer in this series.

There are similarities in duplication of gene which are seen in both TN pure DCIS and TN invasive breast disease and these are not further addressed in this initial analysis.

There are similarities in gains found in both TN pure DCIS and TN invasive breast disease so these are not specifically addressed in this initial analysis. However, there are genomic gains present in some samples of TN invasive breast disease not observed in TN pure DCIS. PANTHER analysis reveals 5 mapped genes. None of these genes have been linked in previous studies to tumourigenesis or breast cancer.

There are no Sc gains in TN pure DCIS in this series. Sc gains are, however, present in TN invasive breast disease. PANTHER analysis reveals 51 mapped genes. This includes the gene ALDH1L1, loss of function or expression of this gene is associated

with decreased apoptosis, increased cell motility, and cancer progression. ALDH1L1 may contribute to ALDH1 activity in breast cancer as high expression of ALDH1A1 mRNA was found to be significantly correlated with poor overall survival in some breast cancer patients (264).

There are similarities found in genomic losses found in TN pure DCIS also present in TN invasive breast disease, these are not examined in this thesis. However, there are genomic losses found in TN invasive breast disease not observed in TN pure DCIS. PANTHER analysis reveals 13 mapped genes. Of these SMAD4 has been suggested as a mediator in the TGF beta pathway acting as a tumour suppressor gene suppressing carcinogenesis (265).

There are total genomic losses present in TN pure DCIS not observed in TN invasive breast disease. PANTHER analysis reveals 41 mapped genes. In this group are 13 olfactory related genes (OLFM4, OR2Z1, OR2Z1, OR4C3, OR4C5, OR4S1, OR4X1, OR4X2, OR7A10, OR7A17, OR7A5, OR7C1, OR7C2). The total loss of OLFM4 in this series is contradictory to previous studies where OLFM4 gene product (Olfactomedin-4) is cited as a candidate biomarker for detection or progression for a variety of solid tumours including breast cancer (266, 267).

There are total genomic losses present in TN invasive breast disease not observed in TN pure DCIS. PANTHER analysis reveals 39 mapped genes. These include the loss of 3 genes encoding for keratins (KRT222, KRT24, KRT25). Loss of keratins has been suggested as one of the possible key components in invasion, although it is acknowledged that loss alone of keratins is not sufficient to cause invasion and/or metastasis (268).

CdLOH found in TN pure DCIS are also present in TN invasive breast disease and have not been analysed in this thesis. There is CdLOH present in TN invasive breast cancer not observed in TN pure DCIS. PANTHER analysis reveals 165 mapped genes. Genes include BCL2, SMAD4, MLH3, alcohol dehydrogenases (ADH1A, ADH1B, ADH4, ADH5, ADH7, ADNP2) and NFKB1. NFKB1 is one of the down regulated genes proposed as part of a seven-gene signature a panel for predicting distant recurrence in patients with triple-negative breast cancers receiving adjuvant chemotherapy following surgery (269).

PANTHER analysis of the CnLOH seen in TN pure DCIS not observed in TN invasive breast disease reveals 335 mapped genes. These include NEDD9 (270) NRD1(271) TGFBR3(272) which have all been associated with invasion in breast cancer. There is CnLOH present in TN invasive breast disease not observed in TN pure DCIS. PANTHER analysis reveals 152 mapped genes. These include BUB3 (273) MCM7 (274) SPIRE1 (261) TGFBR3 (272) which have all been associated with invasion in breast cancer.

5.11.3 Observations of Genomic Analysis of Pure DCIS Compared to Invasive Breast Disease.

See sections 5.8.5 and 5.8.9. As with other series the genomic aberrations for amplifications, duplications, gains, losses, Sc gains and total loss for invasive ER positive breast disease are more pronounced than compared to pure ER positive DCIS. For TN pure DCIS and TN invasive breast disease this is similar without any amplifications or duplications. However, there is greater variation in the chromosomal regions affected for all the pure DCIS compared to the invasive breast disease. This is partially accounted for by the unmatched samples being excluded in this part of the

analysis. However there may be potential genomic regions of interest that identify markers only found in pure DCIS i.e. lesions that are unlikely to progress. Further work is required to elucidate these potential markers.

5.12 Summary of HER2 positive DCIS versus ER positive DCIS and triple negative DCIS

(See section 5.8.11)

Frequency plots for HER2 positive pure DCIS compared to ER positive pure DCIS and TN positive pure DCIS show a surprising similarity. As expected, amplification on chromosome 17 has a much higher frequency than in TN or ER positive pure DCIS as this is the locus of the ERBB2 gene which encodes for the HER2 protein and thus the definition of HER2 positivity in the series. However, CdLOH found on chromosomes 8 and 17, duplication on 8, gains on 3, 8, 17, 20, 21, losses on 6, 8, 17 Sc gains on 3, 8 and total losses on 13, 17 all have similarities between the two groups, albeit at different frequencies.

These findings may indicate that, whilst there are specific genomic changes that are present in immunohistochemical phenotypes that may be suitable targets for therapeutic decisions, these may be independent of genes that are responsible for either the tumourigenesis of DCIS or its progression to an invasive phenotype.

There are differences, as expected, between the two sets particularly with total losses in HER2 positive DCIS found on chromosomes 1, 3, 5, 6, 8, 9, 10, 13, 18, 19, 21 and X seen in the HER2 cohort.

There are similarities between amplifications found in HER2 positive DCIS, ER positive DCIS and TN DCIS. These are not addressed in this initial analysis.

There were duplications present in HER2 positive DCIS not observed in ER and TN DCIS. PANTHER analysis reveals 58 mapped genes. Of these 17 are genes related to the LCE (late cornified envelope) gene cluster within the epidermal differentiation complex on chromosome 1. These have previously been reported in a mouse model of triple negative breast cancer as potential novel driver mutations (275). They have not been previously linked to HER2 positive DCIS.

There are genomic gains in HER2 positive pure DCIS not observed in ER positive and TN pure DCIS. PANTHER analysis reveals 47 mapped genes. Of these genes MACROD2 is has been suggested as a possible cause of Tamoxifen resistance when overexpressed in metastatic tumours (249) but has not been described in association with DCIS.

Genomic gains and Sc gains found in ER positive DCIS, TN DCIS associated with genomic gains in HER2 Pure DCIS demonstrate similar genetic profiles and are not discussed in this work.

There are losses of genes present in HER2 positive DCIS which were not observed in ER and TN positive DCIS. PANTHER analysis reveals 48 mapped genes. One of these CDT1 is an integral part of the pre-replication complex associated with geminin and minichromosome maintenance proteins involved in cellular proliferation (276). Loss of CDT1 expression would be expected to inhibit cellular proliferation and thus may be of importance in preventing further progression of DCIS.

Conversely, there were losses present in ER and TN pure DCIS not observed in HER2 positive pure DCIS. PANTHER analysis identifies 1 gene (MUC16). Increased

serum levels of Ca125 (Muc16) have been shown to correlate with increased pathological grade of breast carcinoma (277).

Total losses in HER2 positive pure DCIS observed in ER and TN pure DCIS show similarity and are not examined in this preliminary research.

There were total losses present in ER pure DCIS and TN pure DCIS not observed in HER2 positive pure DCIS. PANTHER analysis identifies 39 mapped genes. As previously noted (section 5.4.3.2) there are several olfactory genes (OLFM4, OR1A1, OR1A2, OR1D2, OR1E1, OR1E2, OR1G1, OR2Z1, OR2Z1, OR3A1, OR3A2, OR3A3, OR6X1, OR8D1, OR8D2) showing total loss within TN DCIS.

There is CdLOH present in HER2 positive pure DCIS not observed in ER and TN pure DCIS. PANTHER analysis reveals 50 mapped genes. Of these one, IKBKB, was used in a panel of 21 breast cancer related genes compared in DCIS and associated invasive breast disease. The authors state there “were no significant differences in copy number DCIS and adjacent IDC, indicating that DCIS is genetically as advanced as its invasive counterpart” (278).

In the CdLOH analysis comparing ER positive and TN pure DCIS with HER2 positive pure DCIS, one gene (Muc6) is been identified (see total loss in this paragraph).

.

There is CnLOH present in HER2 DCIS not observed in ER positive and TN DCIS. PANTHER analysis identifies 30 mapped genes. DDX3 is an RNA helicase that has antiapoptotic properties, and promotes proliferation and transformation (279). It has been suggested that the protein product could be a potential therapeutic target as overexpression demonstrated by IHC may indicate an oncogenic role (279).

There is CnLOH present in ER and TN pure DCIS not observed in HER2 positive pure DCIS. PANTHER analysis identifies 292 mapped genes. There are several interleukin genes identified (IL12RB2, IL17A, IL23R, ILKAP). Interleukins are key regulators of immune response which may influence increased risk of developing breast cancer (280).

5.12.1 Observations of Genomic Analysis of Pure Her2 DCIS compared to Pure ER positive DCIS and Pure TN DCIS

The frequency plots for pure Her2 DCIS show, unsurprisingly, amplification on chromosome 17 not present in ER positive or TN pure DCIS. However, with the exception of cnLOH the severity of genomic changes (amplification, duplication, gains, Sc gains, losses, total losses and cdLOH) is more marked in HER2 pure DCIS than in either pure ER positive or pure TN DCIS. There is a significant degree of homology in chromosomal regions. This homology could indicate that regardless of subtype that all pure DCIS has common genomic aberrations that either initiate and/or maintain the formation of DCIS. Common gains and losses on the same area e.g. chromosomes 3 and 8 would therefore be candidates for further study. One area of large variation within this series is the number of total losses found in pure Her2 DCIS when compared to other types of pure DCIS. The reason for this comparatively high genomic instability remains elusive.

Chapter 6. Conclusions

6.1 Tissue Microarrays

Tissue microarrays have become an established method for analysing tissue samples from archival formalin fixed paraffin wax embedded samples. In this series using traditional tinctorial staining, immunohistochemistry and in situ hybridisation we have shown the versatility of such a technique. TMAs have increased the number of samples that can be analysed using a variety of biomarkers with a relatively little cost. We have demonstrated that some morphology can be retained with larger cores. However, it is also shown that care and planning is required in using TMAs analysis. Design and planning of TMAs should reflect the scientific question being asked but also enable subsequent further studies on the precious tissue resource. Problems associated with scoring methods for immunohistochemistry and the need for standardisation of TMA research have been highlighted in a publication related to this theses (281). In this series of DCIS samples we have analysed a selection of biomarkers with varying success. This DCIS cohort remains as a TMA resource for further studies.

5.13 6.2 Immunohistochemical Studies of DCIS

The first aim of this thesis was “To determine the expression of a range of immunohistochemical markers in DCIS lesions and determine if these markers can identify molecular subtypes similar to those found in invasive breast cancer.”

These immunohistochemical studies of DCIS have demonstrated that DCIS can be classified into different phenotypic subgroups which are essentially similar to those found in invasive breast cancer. The frequencies within the groups differ slightly from those seen in the literature regarding invasive breast cancer. This could be due to several factors: the cohort used in this study was obtained from a central London hospital. Compared to other regions in the United Kingdom, at least, there may be differences in the patient demographic as London is a particularly ethnically diverse population. The choice of antibodies to define the phenotypic subgroups and the methods used to score markers also varies from some of those used in other studies of invasive breast cancer. The need to standardise immunohistochemical reporting and staining methods needs further work.

In this study we found that although immunohistochemical assessment may be used to define subgroups this was not related to the propensity for a DCIS lesion to progress to an invasive phenotype or to remain in a pure DCIS form in this series, i.e. in terms of recurrence of the disease. This does not necessarily exclude the possibility that panels of IHC biomarkers could be used to determine either invasive potential or risk of recurrence of DCIS. The choice of IHC markers in this study was based upon those already established in invasive breast cancer and either used to predict treatment options in some instances (ER and HER2) or with some evidence of prognostic significance in the literature. It is possible that there is an immunohistochemical marker, or panel of markers, yet to be discovered that could identify the potential to switch from DCIS to an invasive phenotype. It is also possible that there may not be such a marker or panel and the heterogeneity of DCIS prevents such prediction with this 'routine' technique.

6.3 Genomic Studies of DCIS

The second aim of this study was to create a database of significant size providing information on DCIS samples with genomic, immunohistochemical and demographic data. In this respect the aim has been fulfilled. Chapter 5 contains only a small fraction of the data that is available to be analysed and even in this small subset of the available data there is scope for greater analysis. The entire database has great potential for a number of research questions. Should a researcher wish to examine the role of a single gene of interest and its expression in pure DCIS, DCIS associated with invasion or even its absence then a search of the files can be carried out. Alternatively comparison of groups either currently acknowledged such as the basal-like group or groups not yet identified could be elucidated from this data. This study has produced one of the largest genomic data sets on DCIS created to date. It is hoped that this valuable resource will help define the molecular diversity found in both precursor and invasive breast disease to enable improved treatment options for patients.

One advantage of this dataset is the nature of the information provided. This data is not simply a list of genes that may or may not have potential as biomarkers for therapeutic intervention but also contains demographic data on patients, protein expression, immunohistochemical profiles and a number of genomic identifiers. This includes information on copy neutral loss of heterozygosity which is not produced through previous genomic studies. However, the advent of whole genome profiling at reduced costs does indicate that such data sets for comparison will be available in the future building upon the work produced here. The data is currently being analysed by the King's College bioinformatics team and further insights of how this information can be used are expected.

The third aim of this thesis was to undertake an analysis of archival DCIS samples to identify genomic variation, abnormalities and changes between different pure DCIS

lesions and DCIS lesions associated with invasion. This aim has only been partially fulfilled.

Given the vast amount of data available it is not possible within the realm of this thesis to address each of the possible genes that may, or may not, be involved in progression to an invasive state. The copy number aberrations discussed here are a small subset of all the potential markers identified in this heterogeneous disease. The author also acknowledges that the selection of candidate genes here is based upon a combination of selection based upon literature review of breast cancer pathways and personal choice (see 6.4 Limitations of the Study).

There are few studies have examined the genomic diversity of DCIS in a large number of samples. This is one of the largest to date utilising the relatively new molecular inversion probe array platform. Given the known heterogeneity of DCIS, we focussed upon comparing smaller defined sub-groups of DCIS identified using a surrogate immunopanel for the molecular phenotypes as described in invasive cancer by Perou et al (76). Pure DCIS, DCIS associated with invasive disease and invasive disease have all been compared. Whilst chromosomal changes have been shown across all types of DCIS there with similar frequencies, differences have been identified between DCIS subtypes (triple negative, HER2 positive and oestrogen positive). Some of these are expected (i.e. *HER2* gene amplification on chromosome 17) but many others have yet to be fully analysed. Additionally there are also similarities between pure DCIS, DCIS associated with invasion and invasive breast disease. The overall similarity of copy number aberrations in both DCIS and invasive breast cancer indicates that DCIS already has a significant genetic heterogeneity compared to normal mammary epithelium. It may be that the transition from normal tissue to DCIS is where key drivers of tumourigenesis are found. However, further analysis of DCIS at the genomic level has may be required to identify if all key drivers of the invasive phenotype are in place are at the DCIS stage or if further CNVs are required.

Differences in different types of DCIS and their associated invasive tumours have demonstrated a vast array of copy number changes and identified a host of genes some of which may be involved in the transition to an invasive state.

The use of this MIP arrays to analyse over 80 DCIS and associated invasive breast disease samples has resulted in a vast amount of genomic data. It has to be acknowledged that this thesis cannot hope to give a robust and systematic analysis of all of this data. This work has outlined a few of the genes that may be of significance in the transition state from DCIS to invasion. Whilst there are possible flaws in the selection criteria for the genes highlighted (see limitations of the study), this work has provided a significant amount of genomic data previously unavailable in such a large series which will prove invaluable for further studies on DCIS. In addition to the large amount of information available this series of samples can be analysed further within distinctive subgroups; analysis of high grade versus low grade DCIS, pure DCIS versus DCIS associated with invasion and invasive breast disease are available for further study. Unlike many previous studies, the DCIS samples studied here are fully classified in terms of histopathological features, immunoprofiling and a large amount of detailed genomic data and therefore can be segregated into a variety of subgroups to help elucidate the changes that give rise to such heterogeneity in DCIS. It is hoped that further work upon this series of genomic data will continue to help further our understanding of DCIS and its subsequent progression to invasive disease thus facilitating the development of better clinical treatments and decisions.

The final aim of this study was to identify any protein and/or genomic changes in all DCIS (pure and that associated with invasion) compared to invasive breast disease to determine potential biomarkers responsible for progression to an invasive state. In this respect the study has fallen short of the stated aim. Identification of specific

proteins and genes still remains elusive. However, this body of work has produced a dataset that could potentially be used in conjunction with other research to help determine such candidates in the future. The author acknowledges that this is a challenging proposition given the diversity shown at the molecular level of both precursor lesions and invasive forms of breast neoplasms.

6.4 Limitations of the Study

In this study, samples of DCIS and associated invasive tissues were examined. Due to the natural progression of breast cancers, the timeframe for collection of these samples had to, by necessity, particularly for the DCIS cases, be over a significant period. The analysis of a large cohort of samples over an extended time period is fraught with certain problems. These include, but may not be exclusive of, the following:

6.4.1 *Availability of clinical data*

Clinical data over a time period may be lost due to changes in saving such data and changes in the fine details and definitions of the dataset collated. The time period for the samples in this study encompassed the switch from a paper-based system and through different computer-based systems and data from one system may have been incompatible with newer or alternative systems. The King's College Breast Cancer Tissue and Data Bank provided a most thorough examination and upgrade of the clinical data available. However, it has to be acknowledged that the data provided to the bank did have omissions. These gaps could be for a variety of reasons; patients moving to different localities, poor data storage, and lack of details in original patient files. The latter is the most likely cause of missing data; as our understanding of the

complexities of breast cancer have improved the information required and subsequent data collection have become increasingly complicated.

As our understanding of breast cancer has improved, so have treatment options. The introduction of new drugs (new chemotherapy agents such as the taxanes and anthracyclines, trastuzumab, tamoxifen and the aromatase inhibitors) and radiotherapy options have varied over time. Additionally, surgical considerations such as the historical use of mastectomy over wide local excision have reduced the number of cases with follow-up samples.

6.4.2 Technical Considerations

Alongside therapeutic and surgical advances, there have been significant changes in the technical and laboratory treatment of samples. Improved fixation, reduced cold ischaemic times, slicing of large samples, better processing and advances in microscopy have all improved the quality of tissue being analysed. This has led to more accurate and better reporting of disease states.

Immunohistochemistry is widely used as a diagnostic tool in histology. Improvements in antibody clones, immunohistochemical detection methods, antigen retrieval and automation have improved the somewhat capricious nature of this science. Accordingly the introduction of external quality assurance schemes to ensure reproducibly and correct laboratory procedures are followed has improved standardisation of this technique. This does however introduce disparities to samples taken over a long time period. Original reporting and scoring of markers, for example, oestrogen receptor have changed as clones have improved and clinical cut-offs have changed. Historically negative ER cases may now be considered ER positive due to enhanced detection. Re-evaluation may not be entirely accurate as changes in the use of fixatives and processing may produce false negative or even false positive

results. Assessment of historical tissue samples should therefore be judged with caution with known positives and negatives from the same time period assessed in the modern setting.

6.4.3 Genomics

The science of genomics has moved at an astonishing pace since the release of the first human genome in 2000 AD. Whole genome analysis is now both available within a relatively short timeframe at much reduced costs. This has huge benefits for research. However this has raised issues regarding the interpretation of such vast amounts of data. This study has undertaken one of the larger genomic studies of DCIS to date using MIP array analysis of over eighty samples. This created a vast amount of bioinformatics data giving almost whole genome wide information on all of these samples. It is beyond the scope of a single PhD thesis to analyse such data, so a compromise on the best approach had to be taken. In this study comparison of different types of DCIS was first undertaken. This created gene lists of often of thousands of possible genes to analyse for each particular genomic aberration (amplification, duplication, CdLOH, CnLOH, gains, losses, total losses, Sc gains). It was after discussion with the bioinformatics team decided appropriate to concentrate on those samples that showed the greatest variation between each other i.e. where copy number differences occurred in one set of samples (e.g. ER positive pure DCIS) but not another (e.g. TN pure DCIS). This gave a slightly more manageable subset ranging from an estimated 1% to 10% of data to be analysed, dependent upon the particular CNA comparison. Even with this reduction, there were still hundreds of possible genes to analyse. Those suggested here have no basis in selection other than they have been cited previously in the literature. This does raise the possibility of exclusion based upon literary bias. In defence of the selection criteria is the understanding at the outset that that this work was never destined to be a stand-alone

study and that further analysis would be essential. This thesis may therefore be regarded as an initial analysis.

6.5 Further Work

As mentioned to in the previous sections, there is a significant amount of further work that may be undertaken beyond this study. For the immunohistochemical analysis there are many antibody markers that could be assessed in the remaining TMA series. In the selection for this study, certain markers of interest in apoptosis (BCL2, TUNEL, survivin), stem cell markers (ALDH1, CD44, CD24), proliferation markers (geminin) were all omitted based solely on considerations of time and cost. These are just a few of the hundreds of markers already available, not including those in development. One of the advantages of the TMA construction in this series is that there is a reserve of tissue that has already has pathological and genomic analysis performed. This will hopefully help define the suitability of this cohort for use with a particular marker of interest.

The genomic component of this thesis has yielded a comprehensive resource for over eighty samples of DCIS lesions. The analysis of these data sets has only just begun. In the present study only a fraction of the data has been examined with a limited number of genes identified. The data mining methods here will be evaluated for their suitability in determining progression from pure DCIS to an invasive phenotype as well as differences between various forms of DCIS lesion. Different methods of analysis can be explored and further extraction of DNA from samples will provide an additional resource for future studies. As molecular techniques become further enhanced these data could be used as a comparative set to validate further methods.

References

1. F.P. OM, S.E. P, A.M M. Breast Pathology. 2nd ed. R GJ, editor: Elsevier Saunders; 2011.
2. Kuhl CK, Schrading S, Bieling HB, Wardelmann E, Leutner CC, Koenig R, et al. MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. *Lancet*. 2007;370(9586):485-92.
3. Evans.A, Pinder.S.E, Puroshotham.A, Bundred.N. Uncertainties in the management of screen-detected ductal carcinoma in situ. In: Programs NCS, editor. 1st ed. Sheffield: NHSBSP; July 2008. p. 28.
4. Boughey JC, Gonzalez RJ, Bonner E, Kuerer HM. Current treatment and clinical trial developments for ductal carcinoma in situ of the breast. *Oncologist*. 2007;12(11):1276-87.
5. Kumar V, Abbas K. MBBS N, Nelson F, J A, K. KVA. Robbins & Contran Pathological Basis of Disease: Saunders; 2010. 1464 p.
6. Program NHSBS. NHS Breast Screening Programme Review 2012. NHSBHP: <http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp-annualreview2012.pdf>; 2012.
7. Breast Cancer Statistics <http://www.cancerresearchuk.org/cancer-info/cancerstats/types/breast/2014> [
8. **NIH Consensus and State-of-the-Science Statements**. September 22-24 ed2009.
9. SEER Report Cancer Statistics. In: (NPCR) NPoCR, editor. National Program of Cancer Registries (NPCR)2010.
10. Joslyn SA. Ductal carcinoma in situ: trends in geographic, temporal, and demographic patterns of care and survival. *Breast J*. 2006;12(1):20-7.
11. Pappo I, Wasserman I, Halevy A. Ductal carcinoma in situ of the breast in men: a review. *Clin Breast Cancer*. 2005;6(4):310-4.
12. Smith KL, Adank M, Kauff N, Lafaro K, Boyd J, Lee JB, et al. BRCA mutations in women with ductal carcinoma in situ. *Clin Cancer Res*. 2007;13(14):4306-10.
13. Hwang ES, McLennan JL, Moore DH, Crawford BB, Esserman LJ, Ziegler JL. Ductal carcinoma in situ in BRCA mutation carriers. *J Clin Oncol*. 2007;25(6):642-7.
14. Kerlikowske K, Barclay J, Grady D, Sickles EA, Ernster V. Comparison of risk factors for ductal carcinoma in situ and invasive breast cancer. *Journal of the National Cancer Institute*. 1997;89(1):76-82.
15. Claus EB, Stowe M, Carter D. Breast carcinoma in situ: risk factors and screening patterns. *J Natl Cancer Inst*. 2001;93(23):1811-7.
16. Kabat GC, Kim M, Shikany JM, Rodgers AK, Wactawski-Wende J, Lane D, et al. Alcohol consumption and risk of ductal carcinoma in situ of the breast in a cohort of postmenopausal women. *Cancer Epidemiol Biomarkers Prev*. 2010;19(8):2066-72.
17. Kabat GC, Kim M, Kakani C, Tindle H, Wactawski-Wende J, Ockene JK, et al. Cigarette smoking in relation to risk of ductal carcinoma in situ of the breast in a cohort of postmenopausal women. *Am J Epidemiol*. 2010;172(5):591-9.
18. Claus EB, Stowe M, Carter D. Oral contraceptives and the risk of ductal breast carcinoma in situ. *Breast Cancer Res Treat*. 2003;81(2):129-36.
19. Phillips LS, Millikan RC, Schroeder JC, Barnholtz-Sloan JS, Levine BJ. Reproductive and hormonal risk factors for ductal carcinoma in situ of the breast. *Cancer Epidemiol Biomarkers Prev*. 2009;18(5):1507-14.
20. The Nurses Health Study http://www.channing.harvard.edu/nhs/?page_id=70 [

21. Collins LC, Tamimi RM, Baer HJ, Connolly JL, Colditz GA, Schnitt SJ. Outcome of patients with ductal carcinoma in situ untreated after diagnostic biopsy: results from the Nurses' Health Study. *Cancer*. 2005;103(9):1778-84.
22. Betsill WL, Rosen PP, Lieberman PH, Robbins GF. Intraductal carcinoma. Long-term follow-up after treatment by biopsy alone. *JAMA*. 1978;239(18):1863-7.
23. Evans A. The diagnosis and management of pre-invasive breast disease: radiological diagnosis. *Breast Cancer Res*. 2003;5(5):250-3.
24. Yamada T, Mori N, Watanabe M, Kimijima I, Okumoto T, Seiji K, et al. Radiologic-pathologic correlation of ductal carcinoma in situ. *Radiographics*. 2010;30(5):1183-98.
25. Cho KR, Seo BK, Kim CH, Whang KW, Kim YH, Kim BH, et al. Non-calcified ductal carcinoma in situ: ultrasound and mammographic findings correlated with histological findings. *Yonsei Med J*. 2008;49(1):103-10.
26. Breast Pathology. 2nd ed. Philadelphia: Elsevier Saunders; 2011.
27. Boetes C, Mann RM. Ductal carcinoma in situ and breast MRI. *Lancet*. 2007;370(9586):459-60.
28. Schnitt SJ. The diagnosis and management of pre-invasive breast disease: flat epithelial atypia--classification, pathologic features and clinical significance. *Breast Cancer Res*. 2003;5(5):263-8.
29. O'Malley FP, Mohsin SK, Badve S, Bose S, Collins LC, Ennis M, et al. Interobserver reproducibility in the diagnosis of flat epithelial atypia of the breast. *Mod Pathol*. 2006;19(2):172-9.
30. Bombonati A, Sgroi DC. The molecular pathology of breast cancer progression. *J Pathol*. 2011;223(2):307-17.
31. Lopez-Garcia MA, Geyer FC, Lacroix-Triki M, Marchi  C, Reis-Filho JS. Breast cancer precursors revisited: molecular features and progression pathways. *Histopathology*. 2010;57(2):171-92.
32. Pathology Reporting Of Breast Disease. NHSBSP Publication Number 58 January 2005. NHS Cancer Screening Proammes. Report No.: Contract No.
33. Quinn CM, Ostrowski JL. Cytological and architectural heterogeneity in ductal carcinoma in situ of the breast. *J Clin Pathol*. 1997;50(7):596-9.
34. C FR. Pre-Invasive Disease: Pathogenesis and Clinical Management. 1st ed: Springer; 2011. 520 p.
35. Holland R, Hendriks JH, Vebeek AL, Mravunac M, Schuurmans Stekhoven JH. Extent, distribution, and mammographic/histological correlations of breast ductal carcinoma in situ. *Lancet*. 1990;335(8688):519-22.
36. Mai KT, Yazdi HM, Burns BF, Perkins DG. Pattern of distribution of intraductal and infiltrating ductal carcinoma: a three-dimensional study using serial coronal giant sections of the breast. *Hum Pathol*. 2000;31(4):464-74.
37. Thomson JZ, Evans AJ, Pinder SE, Burrell HC, Wilson AR, Ellis IO. Growth pattern of ductal carcinoma in situ (DCIS): a retrospective analysis based on mammographic findings. *Br J Cancer*. 2001;85(2):225-7.
38. Silverstein MJ, Poller DN, Waisman JR, Colburn WJ, Barth A, Gierson ED, et al. Prognostic classification of breast ductal carcinoma-in-situ. *Lancet*. 1995;345(8958):1154-7.
39. Silverstein MJ, Lagios MD, Craig PH, Waisman JR, Lewinsky BS, Colburn WJ, et al. A prognostic index for ductal carcinoma in situ of the breast. *Cancer*. 1996;77(11):2267-74.
40. Tavassoli FA. Ductal carcinoma in situ: introduction of the concept of ductal intraepithelial neoplasia. *Mod Pathol*. 1998;11(2):140-54.
41. Holland R, Peterse JL, Millis RR, Eusebi V, Faverly D, van de Vijver MJ, et al. Ductal carcinoma in situ: a proposal for a new classification. *Semin Diagn Pathol*. 1994;11(3):167-80.

42. Sloane JP, Amendoeira I, Apostolikas N, Bellocq JP, Bianchi S, Boecker W, et al. Consistency achieved by 23 European pathologists in categorizing ductal carcinoma in situ of the breast using five classifications. European Commission Working Group on Breast Screening Pathology. *Hum Pathol.* 1998;29(10):1056-62.
43. Poller DN, Silverstein MJ, Galea M, Locker AP, Elston CW, Blamey RW, et al. Ideas in pathology. Ductal carcinoma in situ of the breast: a proposal for a new simplified histological classification association between cellular proliferation and c-erbB-2 protein expression. *Mod Pathol.* 1994;7(2):257-62.
44. Silverstein MJ. The University of Southern California/Van Nuys prognostic index for ductal carcinoma in situ of the breast. *Am J Surg.* 2003;186(4):337-43.
45. Shoker BS, Sloane JP. DCIS grading schemes and clinical implications. *Histopathology.* 1999;35(5):393-400.
46. Pinder SE, Duggan C, Ellis IO, Cuzick J, Forbes JF, Bishop H, et al. A new pathological system for grading DCIS with improved prediction of local recurrence: results from the UKCCCR/ANZ DCIS trial. *Br J Cancer.* 2010;103(1):94-100.
47. Moran MS, Bai HX, Harris EE, Arthur DW, Bailey L, Bellon JR, et al. ACR Appropriateness Criteria(®) Ductal Carcinoma in Situ. *Breast J.* 2012;18(1):8-15.
48. Donker M, Litière S, Werutsky G, Julien JP, Fentiman IS, Agresti R, et al. Breast-Conserving Treatment With or Without Radiotherapy in Ductal Carcinoma In Situ: 15-Year Recurrence Rates and Outcome After a Recurrence, From the EORTC 10853 Randomized Phase III Trial. *J Clin Oncol.* 2013;31(32):4054-9.
49. Holmberg L, Garmo H, Granstrand B, Ringberg A, Arnesson LG, Sandelin K, et al. Absolute risk reductions for local recurrence after postoperative radiotherapy after sector resection for ductal carcinoma in situ of the breast. *J Clin Oncol.* 2008;26(8):1247-52.
50. Cuzick J, Sestak I, Pinder SE, Ellis IO, Forsyth S, Bundred NJ, et al. Effect of tamoxifen and radiotherapy in women with locally excised ductal carcinoma in situ: long-term results from the UK/ANZ DCIS trial. *Lancet Oncol.* 2011;12(1):21-9.
51. Julien JP, Bijker N, Fentiman IS, Peterse JL, Delledonne V, Rouanet P, et al. Radiotherapy in breast-conserving treatment for ductal carcinoma in situ: first results of the EORTC randomised phase III trial 10853. EORTC Breast Cancer Cooperative Group and EORTC Radiotherapy Group. *Lancet.* 2000;355(9203):528-33.
52. Bijker N, Meijnen P, Peterse JL, Bogaerts J, Van Hoorebeeck I, Julien JP, et al. Breast-conserving treatment with or without radiotherapy in ductal carcinoma-in-situ: ten-year results of European Organisation for Research and Treatment of Cancer randomized phase III trial 10853--a study by the EORTC Breast Cancer Cooperative Group and EORTC Radiotherapy Group. *J Clin Oncol.* 2006;24(21):3381-7.
53. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JM, Brookes C, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. *Eur J Cancer.* 2015;51(16):2296-303.
54. Staley H, McCallum I, Bruce J. Postoperative tamoxifen for ductal carcinoma in situ. *Cochrane Database Syst Rev.* 2012;10:CD007847.
55. Fisher B, Dignam J, Wolmark N, Wickerham DL, Fisher ER, Mamounas E, et al. Tamoxifen in treatment of intraductal breast cancer: National Surgical Adjuvant Breast and Bowel Project B-24 randomised controlled trial. *Lancet.* 1999;353(9169):1993-2000.
56. Nath ME, Robinson TM, Tobon H, Chough DM, Sumkin JH. Automated large-core needle biopsy of surgically removed breast lesions: comparison of samples obtained with 14-, 16-, and 18-gauge needles. *Radiology.* 1995;197(3):739-42.
57. Helbich TH, Rudas M, Haitel A, Kohlberger PD, Thurnher M, Gnant M, et al. Evaluation of needle size for breast biopsy: comparison of 14-, 16-, and 18-gauge biopsy needles. *AJR Am J Roentgenol.* 1998;171(1):59-63.

58. Houssami N, Ciatto S, Ellis I, Ambrogetti D. Underestimation of malignancy of breast core-needle biopsy: concepts and precise overall and category-specific estimates. *Cancer*. 2007;109(3):487-95.
59. Shah VI, Raju U, Chitale D, Deshpande V, Gregory N, Strand V. False-negative core needle biopsies of the breast: an analysis of clinical, radiologic, and pathologic findings in 27 consecutive cases of missed breast cancer. *Cancer*. 2003;97(8):1824-31.
60. Londero V, Zuiani C, Furlan A, Nori J, Bazzocchi M. Role of ultrasound and sonographically guided core biopsy in the diagnostic evaluation of ductal carcinoma in situ (DCIS) of the breast. *Radiol Med*. 2007;112(6):863-76.
61. Rutstein LA, Johnson RR, Poller WR, Dabbs D, Groblewski J, Rakitt T, et al. Predictors of residual invasive disease after core needle biopsy diagnosis of ductal carcinoma in situ. *Breast J*. 2007;13(3):251-7.
62. Chan MY, Lim S. Predictors of invasive breast cancer in ductal carcinoma in situ initially diagnosed by core biopsy. *Asian J Surg*. 2010;33(2):76-82.
63. Hahn M, Kagan KO, Siegmann KC, Krainick-Strobel U, Kraemer B, Fehm T, et al. Mammotome versus ATEC: a comparison of two breast vacuum biopsy techniques under sonographic guidance. *Arch Gynecol Obstet*. 2010;281(2):287-92.
64. Philpotts LE, Shaheen NA, Carter D, Lange RC, Lee CH. Comparison of rebiopsy rates after stereotactic core needle biopsy of the breast with 11-gauge vacuum suction probe versus 14-gauge needle and automatic gun. *AJR Am J Roentgenol*. 1999;172(3):683-7.
65. Shi SR, Cote RJ, Taylor CR. Antigen retrieval techniques: current perspectives. *J Histochem Cytochem*. 2001;49(8):931-7.
66. Park K, Han S, Kim HJ, Kim J, Shin E. HER2 status in pure ductal carcinoma in situ and in the intraductal and invasive components of invasive ductal carcinoma determined by fluorescence in situ hybridization and immunohistochemistry. *Histopathology*. 2006;48(6):702-7.
67. Borgquist S, Zhou W, Jirström K, Amini RM, Sollie T, Sørli T, et al. The prognostic role of HER2 expression in ductal breast carcinoma in situ (DCIS); a population-based cohort study. *BMC Cancer*. 2015;15:468.
68. Kononen J, Bubendorf L, Kallioniemi A, Bärklund M, Schraml P, Leighton S, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med*. 1998;4(7):844-7.
69. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921.
70. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science*. 2001;291(5507):1304-51.
71. Tuna M, Knuutila S, Mills GB. Uniparental disomy in cancer. *Trends Mol Med*. 2009;15(3):120-8.
72. Maciejewski JP, Tiu RV, O'Keefe C. Application of array-based whole genome scanning technologies as a cytogenetic tool in haematological malignancies. *Br J Haematol*. 2009;146(5):479-88.
73. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med*. 2010;61:437-55.
74. Mardis ER. A decade's perspective on DNA sequencing technology. *Nature*. 2011;470(7333):198-203.
75. Hardenbol P, Banér J, Jain M, Nilsson M, Namsaraev EA, Karlin-Neumann GA, et al. Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat Biotechnol*. 2003;21(6):673-8.
76. Perou CM, Sørli T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature*. 2000;406(6797):747-52.

77. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*. 2001;98(19):10869-74.
78. Zhao H, Langerød A, Ji Y, Nowels KW, Nesland JM, Tibshirani R, et al. Different gene expression patterns in invasive lobular and ductal carcinomas of the breast. *Mol Biol Cell*. 2004;15(6):2523-36.
79. Cheang MC, van de Rijn M, Nielsen TO. Gene expression profiling of breast cancer. *Annu Rev Pathol*. 2008;3:67-97.
80. Cheang MC, Voduc D, Bajdik C, Leung S, McKinney S, Chia SK, et al. Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. *Clin Cancer Res*. 2008;14(5):1368-76.
81. Livasy CA, Karaca G, Nanda R, Tretiakova MS, Olopade OI, Moore DT, et al. Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. *Mod Pathol*. 2006;19(2):264-71.
82. Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, Hu Z, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res*. 2004;10(16):5367-74.
83. Zhao X, Rødland EA, Sørlie T, Naume B, Langerød A, Frigessi A, et al. Combining gene signatures improves prediction of breast cancer survival. *PLoS One*. 2011;6(3):e17845.
84. Rakha E, Reis-Filho JS. Basal-like breast carcinoma: from expression profiling to routine practice. *Arch Pathol Lab Med*. 2009;133(6):860-8.
85. Livasy CA, Perou CM, Karaca G, Cowan DW, Maia D, Jackson S, et al. Identification of a basal-like subtype of breast ductal carcinoma in situ. *Hum Pathol*. 2007;38(2):197-204.
86. Dabbs DJ, Chivukula M, Carter G, Bhargava R. Basal phenotype of ductal carcinoma in situ: recognition and immunohistologic profile. *Mod Pathol*. 2006;19(11):1506-11.
87. Bryan BB, Schnitt SJ, Collins LC. Ductal carcinoma in situ with basal-like phenotype: a possible precursor to invasive basal-like breast cancer. *Mod Pathol*. 2006;19(5):617-21.
88. Norum JH, Andersen K, Sørlie T. Lessons learned from the intrinsic subtypes of breast cancer in the quest for precision therapy. *Br J Surg*. 2014;101(8):925-38.
89. Xu J, Wu X, Zhou WH, Liu AW, Wu JB, Deng JY, et al. Aurora-A identifies early recurrence and poor prognosis and promises a potential therapeutic target in triple negative breast cancer. *PLoS One*. 2013;8(2):e56919.
90. Yamamoto S, Yamamoto-Ibusuki M, Yamamoto Y, Fujiwara S, Iwase H. A comprehensive analysis of Aurora A; transcript levels are the most reliable in association with proliferation and prognosis in breast cancer. *BMC Cancer*. 2013;13:217.
91. Ali HR, Dawson SJ, Blows FM, Provenzano E, Leung S, Nielsen T, et al. A Ki67/BCL2 index based on immunohistochemistry is highly prognostic in ER-positive breast cancer. *J Pathol*. 2012;226(1):97-107.
92. Callagy GM, Webber MJ, Pharoah PD, Caldas C. Meta-analysis confirms BCL2 is an independent prognostic marker in breast cancer. *BMC Cancer*. 2008;8:153.
93. Dawson SJ, Makretsov N, Blows FM, Driver KE, Provenzano E, Le Quesne J, et al. BCL2 in breast cancer: a favourable prognostic marker across molecular subtypes and independent of adjuvant therapy received. *Br J Cancer*. 2010;103(5):668-75.
94. Pinder SE, Ellis IO. The diagnosis and management of pre-invasive breast disease: ductal carcinoma in situ (DCIS) and atypical ductal hyperplasia (ADH)--current definitions and classification. *Breast Cancer Res*. 2003;5(5):254-7.

95. Buerger H, Simon R, Schäfer KL, Diallo R, Littmann R, Poremba C, et al. Genetic relation of lobular carcinoma in situ, ductal carcinoma in situ, and associated invasive carcinoma of the breast. *Mol Pathol*. 2000;53(3):118-21.
96. Abd El-Rehim DM, Ball G, Pinder SE, Rakha E, Paish C, Robertson JF, et al. High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. *Int J Cancer*. 2005;116(3):340-50.
97. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ, et al. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol*. 2011;22(8):1736-47.
98. Zhou W, Jirstrom K, Amini RM, Fjällskog ML, Sollie T, Lindman H, et al. Molecular subtypes in ductal carcinoma in situ of the breast and their relation to prognosis: a population-based cohort study. *BMC Cancer*. 2013;13:512.
99. Camp RL, Neumeister V, Rimm DL. A decade of tissue microarrays: progress in the discovery and validation of cancer biomarkers. *J Clin Oncol*. 2008;26(34):5630-7.
100. Kyndi M, Sørensen FB, Knudsen H, Overgaard M, Nielsen HM, Andersen J, et al. Tissue microarrays compared with whole sections and biochemical analyses. A subgroup analysis of DBCG 82 b&c. *Acta Oncol*. 2008;47(4):591-9.
101. Zhang D, Salto-Tellez M, Putti TC, Do E, Koay ES. Reliability of tissue microarrays in detecting protein expression and gene amplification in breast cancer. *Mod Pathol*. 2003;16(1):79-84.
102. Ruiz C, Seibt S, Al Kuraya K, Siraj AK, Mirlacher M, Schraml P, et al. Tissue microarrays for comparing molecular features with proliferation activity in breast cancer. *Int J Cancer*. 2006;118(9):2190-4.
103. Graham AD, Faratian D, Rae F, Thomas JS. Tissue microarray technology in the routine assessment of HER-2 status in invasive breast cancer: a prospective study of the use of immunohistochemistry and fluorescence in situ hybridization. *Histopathology*. 2008;52(7):847-55.
104. Bhargava R, Lal P, Chen B. Feasibility of using tissue microarrays for the assessment of HER-2 gene amplification by fluorescence in situ hybridization in breast carcinoma. *Diagn Mol Pathol*. 2004;13(4):213-6.
105. Sapino A, Marchiò C, Senetta R, Castellano I, Macrì L, Cassoni P, et al. Routine assessment of prognostic factors in breast cancer using a multicore tissue microarray procedure. *Virchows Arch*. 2006;449(3):288-96.
106. Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, et al. InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res*. 2012;40(Database issue):D306-12.
107. McDowall J. Estrogen Receptors
https://www.ebi.ac.uk/interpro/potm/2003_4/Page_1.htm2011 [
108. Jones P, Binns D, McMenamin C, McAnulla C, Hunter S. The InterPro BioMart: federated query and web service access to the InterPro Resource. *Database (Oxford)*. 2011;2011:bar033.
109. Kumar R, Thompson EB. The structure of the nuclear hormone receptors. *Steroids*. 1999;64(5):310-9.
110. Kumar R, Zakharov MN, Khan SH, Miki R, Jang H, Toraldo G, et al. The dynamic structure of the estrogen receptor. *J Amino Acids*. 2011;2011:812540.
111. Enmark E, Gustafsson JA. Oestrogen receptors - an overview. *J Intern Med*. 1999;246(2):133-8.
112. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schütz G, Umesono K, et al. The nuclear receptor superfamily: the second decade. *Cell*. 1995;83(6):835-9.

113. Pike AC, Brzozowski AM, Hubbard RE, Bonn T, Thorsell AG, Engström O, et al. Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist. *EMBO J.* 1999;18(17):4608-18.
114. Segars JH, Driggers PH. Estrogen action and cytoplasmic signaling cascades. Part I: membrane-associated signaling complexes. *Trends Endocrinol Metab.* 2002;13(8):349-54.
115. McDevitt MA, Glidewell-Kenney C, Jimenez MA, Ahearn PC, Weiss J, Jameson JL, et al. New insights into the classical and non-classical actions of estrogen: evidence from estrogen receptor knock-out and knock-in mice. *Mol Cell Endocrinol.* 2008;290(1-2):24-30.
116. Anderson E. The role of oestrogen and progesterone receptors in human mammary development and tumorigenesis. *Breast Cancer Res.* 2002;4(5):197-201.
117. Gown AM. Current issues in ER and HER2 testing by IHC in breast cancer. *Mod Pathol.* 2008;21 Suppl 2:S8-S15.
118. Payne SJ, Bowen RL, Jones JL, Wells CA. Predictive markers in breast cancer--the present. *Histopathology.* 2008;52(1):82-90.
119. Pike MC, Spicer DV, Dahmouch L, Press MF. Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. *Epidemiol Rev.* 1993;15(1):17-35.
120. Kerlikowske K, Molinaro AM, Gauthier ML, Berman HK, Waldman F, Bennington J, et al. Biomarker expression and risk of subsequent tumors after initial ductal carcinoma in situ diagnosis. *J Natl Cancer Inst.* 2010;102(9):627-37.
121. Provenzano E, Hopper JL, Giles GG, Marr G, Venter DJ, Armes JE. Biological markers that predict clinical recurrence in ductal carcinoma in situ of the breast. *Eur J Cancer.* 2003;39(5):622-30.
122. Zhou W, Johansson C, Jirstrom K, Ringberg A, Blomqvist C, Amini RM, et al. A Comparison of Tumor Biology in Primary Ductal Carcinoma In Situ Recurring as Invasive Carcinoma versus a New In Situ. *Int J Breast Cancer.* 2013;2013:582134.
123. Pritchard KI. Breast cancer prevention with selective estrogen receptor modulators: a perspective. *Ann N Y Acad Sci.* 2001;949:89-98.
124. Arpino G, Weiss H, Lee AV, Schiff R, De Placido S, Osborne CK, et al. Estrogen receptor-positive, progesterone receptor-negative breast cancer: association with growth factor receptor expression and tamoxifen resistance. *J Natl Cancer Inst.* 2005;97(17):1254-61.
125. Howell A, Cuzick J, Baum M, Buzdar A, Dowsett M, Forbes JF, et al. Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. *Lancet.* 2005;365(9453):60-2.
126. Cuzick J, Sestak I, Baum M, Buzdar A, Howell A, Dowsett M, et al. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. *Lancet Oncol.* 2010;11(12):1135-41.
127. Salomon DS, Brandt R, Ciardiello F, Normanno N. Epidermal growth factor-related peptides and their receptors in human malignancies. *Crit Rev Oncol Hematol.* 1995;19(3):183-232.
128. Moasser MM. The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis. *Oncogene.* 2007;26(45):6469-87.
129. Burgess AW, Cho HS, Eigenbrot C, Ferguson KM, Garrett TP, Leahy DJ, et al. An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. *Mol Cell.* 2003;12(3):541-52.
130. Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell.* 2006;125(6):1137-49.
131. Hubbard SR, Miller WT. Receptor tyrosine kinases: mechanisms of activation and signaling. *Curr Opin Cell Biol.* 2007;19(2):117-23.

132. Arkhipov A, Shan Y, Kim ET, Dror RO, Shaw DE. Her2 activation mechanism reflects evolutionary preservation of asymmetric ectodomain dimers in the human EGFR family. *Elife*. 2013;2:e00708.
133. Eccles SA, Modjtahedi H, Box G, Court W, Sandle J, Dean CJ. Significance of the c-erbB family of receptor tyrosine kinases in metastatic cancer and their potential as targets for immunotherapy. *Invasion Metastasis*. 1994;14(1-6):337-48.
134. Woodburn JR. The epidermal growth factor receptor and its inhibition in cancer therapy. *Pharmacol Ther*. 1999;82(2-3):241-50.
135. Kamath S, Buolamwini JK. Targeting EGFR and HER-2 receptor tyrosine kinases for cancer drug discovery and development. *Med Res Rev*. 2006;26(5):569-94.
136. Flynn JF, Wong C, Wu JM. Anti-EGFR Therapy: Mechanism and Advances in Clinical Efficacy in Breast Cancer. *J Oncol*. 2009;2009:526963.
137. Walker RA, Bartlett JM, Dowsett M, Ellis IO, Hanby AM, Jasani B, et al. HER2 testing in the UK: further update to recommendations. *J Clin Pathol*. 2008;61(7):818-24.
138. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57-70.
139. Scholzen T, Gerdes J. The Ki-67 protein: from the known and the unknown. *J Cell Physiol*. 2000;182(3):311-22.
140. Gerdes J, Schwab U, Lemke H, Stein H. Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. *Int J Cancer*. 1983;31(1):13-20.
141. Brown DC, Gatter KC. Ki67 protein: the immaculate deception? *Histopathology*. 2002;40(1):2-11.
142. Rahmanzadeh R, Hüttmann G, Gerdes J, Scholzen T. Chromophore-assisted light inactivation of pKi-67 leads to inhibition of ribosomal RNA synthesis. *Cell Prolif*. 2007;40(3):422-30.
143. Endl E, Gerdes J. The Ki-67 protein: fascinating forms and an unknown function. *Exp Cell Res*. 2000;257(2):231-7.
144. Schlüter C, Duchrow M, Wohlenberg C, Becker MH, Key G, Flad HD, et al. The cell proliferation-associated antigen of antibody Ki-67: a very large, ubiquitous nuclear protein with numerous repeated elements, representing a new kind of cell cycle-maintaining proteins. *J Cell Biol*. 1993;123(3):513-22.
145. Hofmann K, Bucher P. The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors. *Trends Biochem Sci*. 1995;20(9):347-9.
146. Chelsky D, Ralph R, Jonak G. Sequence requirements for synthetic peptide-mediated translocation to the nucleus. *Mol Cell Biol*. 1989;9(6):2487-92.
147. Dingwall C, Laskey RA. Nuclear targeting sequences--a consensus? *Trends Biochem Sci*. 1991;16(12):478-81.
148. Silver PA. How proteins enter the nucleus. *Cell*. 1991;64(3):489-97.
149. Bullwinkel J, Baron-Lühr B, Lüdemann A, Wohlenberg C, Gerdes J, Scholzen T. Ki-67 protein is associated with ribosomal RNA transcription in quiescent and proliferating cells. *J Cell Physiol*. 2006;206(3):624-35.
150. Darb-Esfahani S, Loibl S, Müller BM, Roller M, Denkert C, Komor M, et al. Identification of biology-based breast cancer types with distinct predictive and prognostic features: role of steroid hormone and HER2 receptor expression in patients treated with neoadjuvant anthracycline/taxane-based chemotherapy. *Breast Cancer Res*. 2009;11(5):R69.
151. Jones RL, Salter J, A'Hern R, Nerurkar A, Parton M, Reis-Filho JS, et al. The prognostic significance of Ki67 before and after neoadjuvant chemotherapy in breast cancer. *Breast Cancer Res Treat*. 2009;116(1):53-68.

152. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst.* 2009;101(10):736-50.
153. Altintas S, Lambein K, Huizing MT, Braems G, Asjoe FT, Hellemans H, et al. Prognostic significance of oncogenic markers in ductal carcinoma in situ of the breast: a clinicopathologic study. *Breast J.* 2009;15(2):120-32.
154. Ruiz Martín J, Pérez Sánchez C, de las Heras Camino A. Three molecular classifications surrogate to four immunohistochemical markers in 374 invasive breast carcinomas with long follow-up: which is better? *Pathol Res Pract.* 2013;209(6):337-44.
155. Bochman ML, Schwacha A. The Mcm complex: unwinding the mechanism of a replicative helicase. *Microbiol Mol Biol Rev.* 2009;73(4):652-83.
156. Madine MA, Swietlik M, Pelizon C, Romanowski P, Mills AD, Laskey RA. The roles of the MCM, ORC, and Cdc6 proteins in determining the replication competence of chromatin in quiescent cells. *J Struct Biol.* 2000;129(2-3):198-210.
157. Maine GT, Sinha P, Tye BK. Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics.* 1984;106(3):365-85.
158. Takahashi K, Yamada H, Yanagida M. Fission yeast minichromosome loss mutants mis cause lethal aneuploidy and replication abnormality. *Mol Biol Cell.* 1994;5(10):1145-58.
159. Chong JP, Blow JJ. DNA replication licensing factor. *Prog Cell Cycle Res.* 1996;2:83-90.
160. Bochman ML, Schwacha A. The Mcm2-7 complex has in vitro helicase activity. *Mol Cell.* 2008;31(2):287-93.
161. Lei M, Tye BK. Initiating DNA synthesis: from recruiting to activating the MCM complex. *J Cell Sci.* 2001;114(Pt 8):1447-54.
162. Bell SP, Mitchell J, Leber J, Kobayashi R, Stillman B. The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing. *Cell.* 1995;83(4):563-8.
163. Nishitani H, Lygerou Z, Nishimoto T, Nurse P. The Cdt1 protein is required to license DNA for replication in fission yeast. *Nature.* 2000;404(6778):625-8.
164. Bochman ML, Bell SP, Schwacha A. Subunit organization of Mcm2-7 and the unequal role of active sites in ATP hydrolysis and viability. *Mol Cell Biol.* 2008;28(19):5865-73.
165. Gonzalez MA, Pinder SE, Callagy G, Vowler SL, Morris LS, Bird K, et al. Minichromosome maintenance protein 2 is a strong independent prognostic marker in breast cancer. *J Clin Oncol.* 2003;21(23):4306-13.
166. Sanchez-Berrondo J, Mesa P, Ibarra A, Martínez-Jiménez MI, Blanco L, Méndez J, et al. Molecular architecture of a multifunctional MCM complex. *Nucleic Acids Res.* 2012;40(3):1366-80.
167. Freeman A, Morris LS, Mills AD, Stoeber K, Laskey RA, Williams GH, et al. Minichromosome maintenance proteins as biological markers of dysplasia and malignancy. *Clin Cancer Res.* 1999;5(8):2121-32.
168. Sirieix PS, O'Donovan M, Brown J, Save V, Coleman N, Fitzgerald RC. Surface expression of minichromosome maintenance proteins provides a novel method for detecting patients at risk for developing adenocarcinoma in Barrett's esophagus. *Clin Cancer Res.* 2003;9(7):2560-6.
169. Scott IS, Morris LS, Bird K, Davies RJ, Vowler SL, Rushbrook SM, et al. A novel immunohistochemical method to estimate cell-cycle phase distribution in archival tissue: implications for the prediction of outcome in colorectal cancer. *J Pathol.* 2003;201(2):187-97.
170. Shetty A, Loddo M, Fanshawe T, Prevost AT, Sainsbury R, Williams GH, et al. DNA replication licensing and cell cycle kinetics of normal and neoplastic breast. *Br J Cancer.* 2005;93(11):1295-300.

171. Gonzalez MA, Pinder SE, Callagy G, Vowler SL, Morris LS, Bird K, et al. Minichromosome maintenance protein 2 is a strong independent prognostic marker in breast cancer. *J Clin Oncol.* 2003;21(23):4306-13.
172. Gonzalez MA, Tachibana KE, Chin SF, Callagy G, Madine MA, Vowler SL, et al. Geminin predicts adverse clinical outcome in breast cancer by reflecting cell-cycle progression. *J Pathol.* 2004;204(2):121-30.
173. Wojnar A, Kobierzycki C, Krolicka A, Pula B, Podhorska-Okolow M, Dziegiel P. Correlation of Ki-67 and MCM-2 proliferative marker expression with grade of histological malignancy (G) in ductal breast cancers. *Folia Histochem Cytobiol.* 2010;48(3):442-6.
174. Saydam O, Senol O, Schaaïj-Visser TB, Pham TV, Piersma SR, Stemmer-Rachamimov AO, et al. Comparative protein profiling reveals minichromosome maintenance (MCM) proteins as novel potential tumor markers for meningiomas. *J Proteome Res.* 2010;9(1):485-94.
175. Honeycutt KA, Chen Z, Koster MI, Miers M, Nuchtern J, Hicks J, et al. Deregulated minichromosomal maintenance protein MCM7 contributes to oncogene driven tumorigenesis. *Oncogene.* 2006;25(29):4027-32.
176. Williams GH, Swinn R, Prevost AT, De Clive-Lowe P, Halsall I, Going JJ, et al. Diagnosis of oesophageal cancer by detection of minichromosome maintenance 5 protein in gastric aspirates. *Br J Cancer.* 2004;91(4):714-9.
177. Wharton SB, Chan KK, Anderson JR, Stoeber K, Williams GH. Replicative Mcm2 protein as a novel proliferation marker in oligodendrogliomas and its relationship to Ki67 labelling index, histological grade and prognosis. *Neuropathol Appl Neurobiol.* 2001;27(4):305-13.
178. Meng MV, Grossfeld GD, Williams GH, Dilworth S, Stoeber K, Mulley TW, et al. Minichromosome maintenance protein 2 expression in prostate: characterization and association with outcome after therapy for cancer. *Clin Cancer Res.* 2001;7(9):2712-8.
179. Reena RM, Mastura M, Siti-Aishah MA, Munirah MA, Norlia A, Naqiyah I, et al. Minichromosome maintenance protein 2 is a reliable proliferative marker in breast carcinoma. *Ann Diagn Pathol.* 2008;12(5):340-3.
180. Nasir A, Chen DT, Gruidl M, Henderson-Jackson EB, Venkataramu C, McCarthy SM, et al. Novel molecular markers of malignancy in histologically normal and benign breast. *Patholog Res Int.* 2011;2011:489064.
181. Taylor-Papadimitriou J, Stampfer M, Bartek J, Lewis A, Boshell M, Lane EB, et al. Keratin expression in human mammary epithelial cells cultured from normal and malignant tissue: relation to in vivo phenotypes and influence of medium. *J Cell Sci.* 1989;94 (Pt 3):403-13.
182. Moll R, Franke WW, Schiller DL, Geiger B, Krepler R. The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *Cell.* 1982;31(1):11-24.
183. Jarasch ED, Nagle RB, Kaufmann M, Maurer C, Böcker WJ. Differential diagnosis of benign epithelial proliferations and carcinomas of the breast using antibodies to cytokeratins. *Hum Pathol.* 1988;19(3):276-89.
184. Cleator S, Heller W, Coombes RC. Triple-negative breast cancer: therapeutic options. *Lancet Oncol.* 2007;8(3):235-44.
185. Ross JS, Fletcher JA. The HER-2/neu Oncogene in Breast Cancer: Prognostic Factor, Predictive Factor, and Target for Therapy. *Oncologist.* 1998;3(4):237-52.
186. Zhao J, Wu R, Au A, Marquez A, Yu Y, Shi Z. Determination of HER2 gene amplification by chromogenic in situ hybridization (CISH) in archival breast carcinoma. *Mod Pathol.* 2002;15(6):657-65.

187. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst.* 2011;103(22):1656-64.
188. Walker RA, Hanby A, Pinder SE, Thomas J, Ellis IO, Subgroup NCCfBPR. Current issues in diagnostic breast pathology. *J Clin Pathol.* 2012;65(9):771-85.
189. Poller DN, Snead DR, Roberts EC, Galea M, Bell JA, Gilmour A, et al. Oestrogen receptor expression in ductal carcinoma in situ of the breast: relationship to flow cytometric analysis of DNA and expression of the c-erbB-2 oncoprotein. *Br J Cancer.* 1993;68(1):156-61.
190. Thomas J, Hanby A, Pinder S, Ellis I, Macartney J, Clements K, et al. Implications of inconsistent measurement of ER status in non-invasive breast cancer: a study of 1,684 cases from the Sloane Project. *Breast J.* 2008;14(1):33-8.
191. Karayiannakis AJ, Bastounis EA, Chatzigianni EB, Makri GG, Alexiou D, Karamanakis P. Immunohistochemical detection of oestrogen receptors in ductal carcinoma in situ of the breast. *Eur J Surg Oncol.* 1996;22(6):578-82.
192. Hannemann J, Velds A, Halfwerk JB, Kreike B, Peterse JL, van de Vijver MJ. Classification of ductal carcinoma in situ by gene expression profiling. *Breast Cancer Res.* 2006;8(5):R61.
193. Thomas J, Evans A, Macartney J, Pinder SE, Hanby A, Ellis I, et al. Radiological and pathological size estimations of pure ductal carcinoma in situ of the breast, specimen handling and the influence on the success of breast conservation surgery: a review of 2564 cases from the Sloane Project. *Br J Cancer.* 2010;102(2):285-93.
194. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, et al. The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics.* 2006;7:96.
195. Sørlie T. Molecular portraits of breast cancer: tumour subtypes as distinct disease entities. *Eur J Cancer.* 2004;40(18):2667-75.
196. Meijnen P, Peterse JL, Antonini N, Rutgers EJ, van de Vijver MJ. Immunohistochemical categorisation of ductal carcinoma in situ of the breast. *Br J Cancer.* 2008;98(1):137-42.
197. Clark SE, Warwick J, Carpenter R, Bowen RL, Duffy SW, Jones JL. Molecular subtyping of DCIS: heterogeneity of breast cancer reflected in pre-invasive disease. *Br J Cancer.* 2011;104(1):120-7.
198. Zhou W, Jirström K, Johansson C, Amini RM, Blomqvist C, Agbaje O, et al. Long-term survival of women with basal-like ductal carcinoma in situ of the breast: a population-based cohort study. *BMC Cancer.* 2010;10:653.
199. Steinman S, Wang J, Bourne P, Yang Q, Tang P. Expression of cytokeratin markers, ER-alpha, PR, HER-2/neu, and EGFR in pure ductal carcinoma in situ (DCIS) and DCIS with co-existing invasive ductal carcinoma (IDC) of the breast. *Ann Clin Lab Sci.* 2007;37(2):127-34.
200. Tamimi RM, Baer HJ, Marotti J, Galan M, Galaburda L, Fu Y, et al. Comparison of molecular phenotypes of ductal carcinoma in situ and invasive breast cancer. *Breast Cancer Res.* 2008;10(4):R67.
201. Tang P, Skinner KA, Hicks DG. Molecular classification of breast carcinomas by immunohistochemical analysis: are we ready? *Diagn Mol Pathol.* 2009;18(3):125-32.
202. Hernandez L, Wilkerson PM, Lambros MB, Campion-Flora A, Rodrigues DN, Gauthier A, et al. Genomic and mutational profiling of ductal carcinomas in situ and matched adjacent invasive breast cancers reveals intra-tumour genetic heterogeneity and clonal selection. *J Pathol.* 2012;227(1):42-52.
203. Ma XJ, Salunga R, Tuggle JT, Gaudet J, Enright E, McQuary P, et al. Gene expression profiles of human breast cancer progression. *Proc Natl Acad Sci U S A.* 2003;100(10):5974-9.

204. Buerger H, Otterbach F, Simon R, Poremba C, Diallo R, Decker T, et al. Comparative genomic hybridization of ductal carcinoma in situ of the breast-evidence of multiple genetic pathways. *J Pathol.* 1999;187(4):396-402.
205. Rydén L, Landberg G, Stål O, Nordenskjöld B, Fernö M, Bendahl PO. HER2 status in hormone receptor positive premenopausal primary breast cancer adds prognostic, but not tamoxifen treatment predictive, information. *Breast Cancer Res Treat.* 2008;109(2):351-7.
206. Alqaisi A, Chen L, Romond E, Chambers M, Stevens M, Pasley G, et al. Impact of estrogen receptor (ER) and human epidermal growth factor receptor-2 (HER2) co-expression on breast cancer disease characteristics: implications for tumor biology and research. *Breast Cancer Res Treat.* 2014;148(2):437-44.
207. Vincent-Salomon A, Lucchesi C, Gruel N, Raynal V, Pierron G, Goudefroye R, et al. Integrated genomic and transcriptomic analysis of ductal carcinoma in situ of the breast. *Clin Cancer Res.* 2008;14(7):1956-65.
208. Yao J, Weremowicz S, Feng B, Gentleman RC, Marks JR, Gelman R, et al. Combined cDNA array comparative genomic hybridization and serial analysis of gene expression analysis of breast tumor progression. *Cancer Res.* 2006;66(8):4065-78.
209. Gao Y, Niu Y, Wang X, Wei L, Lu S. Genetic changes at specific stages of breast cancer progression detected by comparative genomic hybridization. *J Mol Med (Berl).* 2009;87(2):145-52.
210. Liao S, Desouki MM, Gaile DP, Shepherd L, Nowak NJ, Conroy J, et al. Differential copy number aberrations in novel candidate genes associated with progression from in situ to invasive ductal carcinoma of the breast. *Genes Chromosomes Cancer.* 2012;51(12):1067-78.
211. Cowell CF, Weigelt B, Sakr RA, Ng CK, Hicks J, King TA, et al. Progression from ductal carcinoma in situ to invasive breast cancer: revisited. *Mol Oncol.* 2013;7(5):859-69.
212. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. *Nat Genet.* 2004;36(9):949-51.
213. Shlien A, Malkin D. Copy number variations and cancer. *Genome Med.* 2009;1(6):62.
214. R ZN. Copy Number Profiling in Normal and Tumour Genomes 2012.
215. Ueno T, Emi M, Sato H, Ito N, Muta M, Kuroi K, et al. Genome-wide copy number analysis in primary breast cancer. *Expert Opin Ther Targets.* 2012;16 Suppl 1:S31-5.
216. Byler S, Goldgar S, Heerboth S, Leary M, Housman G, Moulton K, et al. Genetic and epigenetic aspects of breast cancer progression and therapy. *Anticancer Res.* 2014;34(3):1071-7.
217. Xu H, Eirew P, Mullaly SC, Aparicio S. The omics of triple-negative breast cancers. *Clin Chem.* 2014;60(1):122-33.
218. Reis-Filho JS, Simpson PT, Gale T, Lakhani SR. The molecular genetics of breast cancer: the contribution of comparative genomic hybridization. *Pathol Res Pract.* 2005;201(11):713-25.
219. Martin V, Botta F, Zanellato E, Molinari F, Crippa S, Mazzucchelli L, et al. Molecular characterization of EGFR and EGFR-downstream pathways in triple negative breast carcinomas with basal like features. *Histol Histopathol.* 2012;27(6):785-92.
220. Ando Y, Iwase H, Ichihara S, Toyoshima S, Nakamura T, Yamashita H, et al. Loss of heterozygosity and microsatellite instability in ductal carcinoma in situ of the breast. *Cancer Lett.* 2000;156(2):207-14.
221. Johnson CE, Gorringe KL, Thompson ER, Opeskin K, Boyle SE, Wang Y, et al. Identification of copy number alterations associated with the progression of DCIS to invasive ductal carcinoma. *Breast Cancer Res Treat.* 2011.

222. Lee S, Stewart S, Nagtegaal I, Luo J, Wu Y, Colditz G, et al. Differentially Expressed Genes Regulating the Progression of Ductal Carcinoma in Situ to Invasive Breast Cancer. *Cancer Res.* 2012.
223. Carraro DM, Elias EV, Andrade VP. Ductal carcinoma in situ of the breast: morphological and molecular features implicated in progression. *Biosci Rep.* 2013.
224. Vargas AC, Reed AE, Waddell N, Lane A, Reid LE, Smart CE, et al. Gene expression profiling of tumour epithelial and stromal compartments during breast cancer progression. *Breast Cancer Res Treat.* 2012.
225. Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc.* 2013;8(8):1551-66.
226. Seligson DB, Horvath S, Shi T, Yu H, Tze S, Grunstein M, et al. Global histone modification patterns predict risk of prostate cancer recurrence. *Nature.* 2005;435(7046):1262-6.
227. Elsheikh SE, Green AR, Rakha EA, Powe DG, Ahmed RA, Collins HM, et al. Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. *Cancer Res.* 2009;69(9):3802-9.
228. Kadota M, Sato M, Duncan B, Ooshima A, Yang HH, Diaz-Meyer N, et al. Identification of novel gene amplifications in breast cancer and coexistence of gene amplification with an activating mutation of PIK3CA. *Cancer Res.* 2009;69(18):7357-65.
229. Lightfoot HM, Lark A, Livasy CA, Moore DT, Cowan D, Dressler L, et al. Upregulation of focal adhesion kinase (FAK) expression in ductal carcinoma in situ (DCIS) is an early event in breast tumorigenesis. *Breast Cancer Res Treat.* 2004;88(2):109-16.
230. Xie F, Wang Q, Chen Y, Gu Y, Mao H, Zeng W, et al. Costimulatory molecule OX40/OX40L expression in ductal carcinoma in situ and invasive ductal carcinoma of breast: an immunohistochemistry-based pilot study. *Pathol Res Pract.* 2010;206(11):735-9.
231. Howard BA. The role of NRG3 in mammary development. *J Mammary Gland Biol Neoplasia.* 2008;13(2):195-203.
232. Kogata N, Oliemuller E, Wansbury O, Howard BA. Neuregulin-3 regulates epithelial progenitor cell positioning and specifies mammary phenotype. *Stem Cells Dev.* 2014;23(22):2758-70.
233. Varley JM, Armour J, Swallow JE, Jeffreys AJ, Ponder BA, T'Ang A, et al. The retinoblastoma gene is frequently altered leading to loss of expression in primary breast tumours. *Oncogene.* 1989;4(6):725-9.
234. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature.* 2012;486(7403):346-52.
235. Visse R, Nagase H. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res.* 2003;92(8):827-39.
236. Cao YW, Wan GX, Zhao CX, Hu JM, Li L, Liang WH, et al. Notch1 single nucleotide polymorphism rs3124591 is associated with the risk of development of invasive ductal breast carcinoma in a Chinese population. *Int J Clin Exp Pathol.* 2014;7(7):4286-94.
237. Yu Q, Li Y, Mu K, Li Z, Meng Q, Wu X, et al. Amplification of Mdmx and overexpression of MDM2 contribute to mammary carcinogenesis by substituting for p53 mutations. *Diagn Pathol.* 2014;9:71.
238. Carlsson H, Petersson S, Enerbäck C. Cluster analysis of S100 gene expression and genes correlating to psoriasin (S100A7) expression at different stages of breast cancer development. *Int J Oncol.* 2005;27(6):1473-81.

239. Moon HG, Hwang KT, Kim JA, Kim HS, Lee MJ, Jung EM, et al. NFIB is a potential target for estrogen receptor-negative breast cancers. *Mol Oncol*. 2011;5(6):538-44.
240. Wetterskog D, Lopez-Garcia MA, Lambros MB, A'Hern R, Geyer FC, Milanezi F, et al. Adenoid cystic carcinomas constitute a genomically distinct subgroup of triple-negative and basal-like breast cancers. *J Pathol*. 2012;226(1):84-96.
241. Soysal S, Obermann EC, Gao F, Oertli D, Gillanders WE, Viehl CT, et al. PTP1B expression is an independent positive prognostic factor in human breast cancer. *Breast Cancer Res Treat*. 2013;137(2):637-44.
242. Fontanil T, Rúa S, Llamazares M, Moncada-Pazos A, Quirós PM, García-Suárez O, et al. Interaction between the ADAMTS-12 metalloprotease and fibulin-2 induces tumor-suppressive effects in breast cancer cells. *Oncotarget*. 2014;5(5):1253-64.
243. Naik MU, Pham NT, Beebe K, Dai W, Naik UP. Calcium-dependent inhibition of polo-like kinase 3 activity by CIB1 in breast cancer cells. *Int J Cancer*. 2011;128(3):587-96.
244. Fonseca-Sánchez MA, Pérez-Plasencia C, Fernández-Retana J, Arechaga-Ocampo E, Marchat LA, Rodríguez-Cuevas S, et al. microRNA-18b is upregulated in breast cancer and modulates genes involved in cell migration. *Oncol Rep*. 2013;30(5):2399-410.
245. Davis LM, Harris C, Tang L, Doherty P, Hrabec P, Sakai Y, et al. Amplification patterns of three genomic regions predict distant recurrence in breast carcinoma. *J Mol Diagn*. 2007;9(3):327-36.
246. Stebbing J, Zhang H, Xu Y, Lit LC, Green AR, Grothey A, et al. KSR1 regulates BRCA1 degradation and inhibits breast cancer growth. *Oncogene*. 2014;0.
247. Rakha EA, Boyce RW, Abd El-Rehim D, Kurien T, Green AR, Paish EC, et al. Expression of mucins (MUC1, MUC2, MUC3, MUC4, MUC5AC and MUC6) and their prognostic significance in human breast cancer. *Mod Pathol*. 2005;18(10):1295-304.
248. Do SI, Kim K, Kim DH, Chae SW, Park YL, Park CH, et al. Associations between the Expression of Mucins (MUC1, MUC2, MUC5AC, and MUC6) and Clinicopathologic Parameters of Human Breast Ductal Carcinomas. *J Breast Cancer*. 2013;16(2):152-8.
249. Mohseni M, Cidado J, Croessmann S, Cravero K, Cimino-Mathews A, Wong HY, et al. MACROD2 overexpression mediates estrogen independent growth and tamoxifen resistance in breast cancers. *Proc Natl Acad Sci U S A*. 2014;111(49):17606-11.
250. Fu A, Leaderer D, Zheng T, Hoffman AE, Stevens RG, Zhu Y. Genetic and epigenetic associations of circadian gene TIMELESS and breast cancer risk. *Mol Carcinog*. 2012;51(12):923-9.
251. Johnson CE, Gorringe KL, Thompson ER, Opeskin K, Boyle SE, Wang Y, et al. Identification of copy number alterations associated with the progression of DCIS to invasive ductal carcinoma. *Breast Cancer Res Treat*. 2012;133(3):889-98.
252. Sun PH, Ye L, Mason MD, Jiang WG. Protein tyrosine phosphatase kappa (PTPRK) is a negative regulator of adhesion and invasion of breast cancer cells, and associates with poor prognosis of breast cancer. *J Cancer Res Clin Oncol*. 2013;139(7):1129-39.
253. Ataseven B, Gunesch A, Eiermann W, Kates RE, Högel B, Knyazev P, et al. PTK7 as a potential prognostic and predictive marker of response to adjuvant chemotherapy in breast cancer patients, and resistance to anthracycline drugs. *Onco Targets Ther*. 2014;7:1723-31.
254. Zhao YF, Zhao JY, Yue H, Hu KS, Shen H, Guo ZG, et al. FOXD1 promotes breast cancer proliferation and chemotherapeutic drug resistance by targeting p27. *Biochem Biophys Res Commun*. 2015;456(1):232-7.

255. Kren BT, Unger GM, Abedin MJ, Vogel RI, Henzler CM, Ahmed K, et al. Preclinical evaluation of cyclin dependent kinase 11 and casein kinase 2 survival kinases as RNA interference targets for triple negative breast cancer therapy. *Breast Cancer Res.* 2015;17(1):19.
256. Mahmoodi M, Nahvi H, Mahmoudi M, Kasaian A, Mohagheghi MA, Divsalar K, et al. HLA-DRB1,-DQA1 and -DQB1 allele and haplotype frequencies in female patients with early onset breast cancer. *Pathol Oncol Res.* 2012;18(1):49-55.
257. Purrington KS, Slettedahl S, Bolla MK, Michailidou K, Czene K, Nevanlinna H, et al. Genetic variation in mitotic regulatory pathway genes is associated with breast tumor grade. *Hum Mol Genet.* 2014;23(22):6034-46.
258. Marchione R, Leibovitch SA, Lenormand JL. The translational factor eIF3f: the ambivalent eIF3 subunit. *Cell Mol Life Sci.* 2013;70(19):3603-16.
259. Silvera D, Formenti SC, Schneider RJ. Translational control in cancer. *Nat Rev Cancer.* 2010;10(4):254-66.
260. Nupponen NN, Porkka K, Kakkola L, Tanner M, Persson K, Borg A, et al. Amplification and overexpression of p40 subunit of eukaryotic translation initiation factor 3 in breast and prostate cancer. *Am J Pathol.* 1999;154(6):1777-83.
261. Lagal V, Abrivard M, Gonzalez V, Perazzi A, Popli S, Verzeroli E, et al. Spire-1 contributes to the invadosome and its associated invasive properties. *J Cell Sci.* 2014;127(Pt 2):328-40.
262. Boudreau HE, Casterline BW, Rada B, Korzeniowska A, Leto TL. Nox4 involvement in TGF-beta and SMAD3-driven induction of the epithelial-to-mesenchymal transition and migration of breast epithelial cells. *Free Radic Biol Med.* 2012;53(7):1489-99.
263. Castellana B, Escuin D, Peiró G, Garcia-Valdecasas B, Vázquez T, Pons C, et al. ASPN and GJB2 Are Implicated in the Mechanisms of Invasion of Ductal Breast Carcinomas. *J Cancer.* 2012;3:175-83.
264. Wu S, Xue W, Huang X, Yu X, Luo M, Huang Y, et al. Distinct prognostic values of ALDH1 isoenzymes in breast cancer. *Tumour Biol.* 2015.
265. Yang G, Yang X. Smad4-mediated TGF-beta signaling in tumorigenesis. *Int J Biol Sci.* 2010;6(1):1-8.
266. Guette C, Valo I, Vétillard A, Coqueret O. Olfactomedin-4 is a candidate biomarker of solid gastric, colorectal, pancreatic, head and neck, and prostate cancers. *Proteomics Clin Appl.* 2015;9(1-2):58-63.
267. Koshida S, Kobayashi D, Moriai R, Tsuji N, Watanabe N. Specific overexpression of OLFM4(GW112/HGC-1) mRNA in colon, breast and lung cancer tissues detected using quantitative analysis. *Cancer Sci.* 2007;98(3):315-20.
268. Loschke F, Seltsmann K, Bouameur JE, Magin TM. Regulation of keratin network organization. *Curr Opin Cell Biol.* 2015;32C:56-64.
269. Park YH, Jung HH, Do IG, Cho EY, Sohn I, Jung SH, et al. A seven-gene signature can predict distant recurrence in patients with triple-negative breast cancers who receive adjuvant chemotherapy following surgery. *Int J Cancer.* 2015;136(8):1976-84.
270. Iida J, Dorchak J, Clancy R, Slavik J, Ellsworth R, Katagiri Y, et al. Role for chondroitin sulfate glycosaminoglycan in NEDD9-mediated breast cancer cell growth. *Exp Cell Res.* 2015;330(2):358-70.
271. Choong LY, Lim SK, Chen Y, Loh MC, Toy W, Wong CY, et al. Elevated NRD1 metalloprotease expression plays a role in breast cancer growth and proliferation. *Genes Chromosomes Cancer.* 2011;50(10):837-47.
272. Elderbroom JL, Huang JJ, Gatza CE, Chen J, How T, Starr M, et al. Ectodomain shedding of TβRIII is required for TβRIII-mediated suppression of TGF-β signaling and breast cancer migration and invasion. *Mol Biol Cell.* 2014;25(16):2320-32.

273. Wang P, Wang Y, Yan H, Xie Q, Zhao L, Xu S, et al. Genetic variation in the major mitotic checkpoint genes and risk of breast cancer: a multigenic study on cancer susceptibility. *Tumour Biol.* 2014;35(7):6701-5.
274. Huang TH, Huo L, Wang YN, Xia W, Wei Y, Chang SS, et al. Epidermal growth factor receptor potentiates MCM7-mediated DNA replication through tyrosine phosphorylation of Lyn kinase in human cancers. *Cancer Cell.* 2013;23(6):796-810.
275. Johnson JP, Kumar P, Koulis M, Patel M, Simin K. Crucial and novel cancer drivers in a mouse model of triple-negative breast cancer. *Cancer Genomics Proteomics.* 2014;11(3):115-26.
276. Gonzalez MA, Tachibana KE, Chin SF, Callagy G, Madine MA, Vowler SL, et al. Geminin predicts adverse clinical outcome in breast cancer by reflecting cell-cycle progression. *J Pathol.* 2004;204(2):121-30.
277. Zhang SJ, Hu Y, Qian HL, Jiao SC, Liu ZF, Tao HT, et al. Expression and significance of ER, PR, VEGF, CA15-3, CA125 and CEA in judging the prognosis of breast cancer. *Asian Pac J Cancer Prev.* 2013;14(6):3937-40.
278. Moelans CB, de Weger RA, Monsuur HN, Maes AH, van Diest PJ. Molecular differences between ductal carcinoma in situ and adjacent invasive breast carcinoma: a multiplex ligation-dependent probe amplification study. *Anal Cell Pathol (Amst).* 2010;33(3):165-73.
279. Bol GM, Raman V, van der Groep P, Vermeulen JF, Patel AH, van der Wall E, et al. Expression of the RNA helicase DDX3 and the hypoxia response in breast cancer. *PLoS One.* 2013;8(5):e63548.
280. Slattery ML, Herrick JS, Torres-Mejia G, John EM, Giuliano AR, Hines LM, et al. Genetic variants in interleukin genes are associated with breast cancer risk and survival in a genetically admixed population: the Breast Cancer Health Disparities Study. *Carcinogenesis.* 2014;35(8):1750-9.
281. Pinder SE, Brown JP, Gillett C, Purdie CA, Speirs V, Thompson AM, et al. The manufacture and assessment of tissue microarrays: suggestions and criteria for analysis, with breast cancer as an example. *J Clin Pathol.* 2012.

Appendices

Appendix 1: Copyright Permissions

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

May 13, 2014

This is a License Agreement between John P Brown ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3386970712499
License date	May 13, 2014
Licensed content publisher	John Wiley and Sons
Licensed content publication	The EMBO Journal
Licensed content title	Cell cycle-dependent regulation of the association between origin recognition proteins and somatic cell chromatin
Licensed copyright line	Copyright © 2002 European Molecular Biology Organization
Licensed content author	Wei-Hsin Sun, Thomas R. Coleman, Melvin L. DePamphilis
Licensed content date	Mar 15, 2002
Start page	1437
End page	1446
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 8
Will you be translating?	No
Title of your thesis / dissertation	Classification of Ductal Carcinoma in Situ of the Human Breast
Expected completion date	Feb 2015
Expected size (number of pages)	250
Total	BP

5.15

5.16 Appendix 2: Excel TMA Maps and IHC Scores

TMA block number	DCIS TMA location	ER Average	PR	HER2	EGFR	CK5/6	CK14	KI67	MCM2
DCIS 15A	A1	0	0	0	0	0	1	1	0
DCIS 15A	A2	0	0	0	0	0	0	x	0
DCIS 15A	A4	2	0	0	0	x	10	1	0
DCIS 15A	A5	0	0	1	3	0	5	40	10
DCIS 15A	B1	x	x	x	0	x	x	x	x
DCIS 15A	B3	0	0	0	3	0	0	10	40
DCIS 15A	B4	0	0	0	0	0	0	30	x
DCIS 15A	B5	2	0	0	1	0	0	15	30
DCIS 15A	C2	5	0	0	0	0	0	5	5
DCIS 15A	C3	1	0	1	2	0	0	5	5
DCIS 15A	C4	5	7	2	0	0	0	1	1
DCIS 15A	C5	8	8	0	0	0	0	0	2
DCIS 15A	D1	x	x	x	x	x	x	x	x
DCIS 15A	D2	x	x	x	x	x	x	x	x
DCIS 15A	D3	0	x	x	x	x	x	x	x
DCIS 15A	D4	x	x	x	x	x	x	x	x
DCIS 15A	E1	2	x	x	x	x	x	x	x

DCIS 15A	E2	0	0	0	0	0	x	0	5
DCIS 15A	E3	5	x	x	x	x	x	x	x
DCIS 15A	E4	8	x	0	0	0	0	1	1
DCIS 16A	A1	0	x	x	x	x	x	x	x
TMA block number	INVASIVE TUMOUR location	ER (Allred)	PR (Allred)	HER2 (HER2)	EGFR (HER2)	CK5/6 (%)	CK14 (%)	KI67 (%)	MCM
DCIS 16A	A2	8	5	x	0	0	0	1	3
DCIS 16A	A4	7	0	3	0	0	0	1	5
DCIS 16A	A5	2	4	2	0	0	0	0	0
DCIS 16A	B1	3	x	x	x	x	x	x	x
DCIS 16A	B3	5	8	2	0	0	0	2	5
DCIS 16A	B4	1	x	x	x	0	0	x	x
DCIS 16A	B5	4	0	0	x	0	0	20	45
DCIS 16A	C2	5	8	2	x	0	0	10	25
DCIS 16A	C3	4	0	1	3	0	0	25	30
DCIS 16A	C4	2	0	0	1	0	0	10	25
DCIS 16A	C5	4	0	3	1	0	0	5	35
DCIS 16A	D1	0	0	0	0	0	0	15	40
DCIS 16A	D2	3	x	x	x	x	x	x	x
DCIS 16A	D3	5	x	x	x	x	x	x	x
DCIS 16A	D4	5	8	0	0	0	0	x	x
DCIS 16A	E1	5	0	x	x	0	0	10	10
DCIS 16A	E2	1	x	3	0	0	0	x	x

DCIS 16A	E3	4	0	3	0	0	0	10	30
DCIS 16A	E4	2	0	2	1	1	0	2	10
DCIS 16A	E5	x	0	x	2	0	0	3	35
INV T 15	A1	0	0	0	3	15	0	8	8
INV T 15	A2	0	0	0	3	0	0	5	2
INV T 15	A4	0	0	0	3	60	95	4	53
INV T 15	A5	0	0	1	3	0	0	50	17
TMA block number	INVASIVE TUMOUR location	ER (Allred)	PR (Allred)	HER2 (HER2)	EGFR (HER2)	CK5/6 (%)	CK14 (%)	KI67 (%)	MCM
INV T 15	A6	0	0	2	3	0	20	30	17
INV T 15	A7	0	0	0	3	0	100	10	37
INV T 15	B1	0	0	0	0	0	0	50	17
INV T 15	B3	4	0	x	2	0	25	20	15
INV T 15	B4	6	0	1	2	0	0	20	7
INV T 15	B5	8	6	0	0	0	0	3	1
INV T 15	B6	8	7	2	0	0	0	15	5
INV T 15	B7	8	8	0	0	0	0	0	0
INV T 15	C2	8	7	1	0	0	0	15	5
INV T 15	C3	8	8	1	0	0	0	4	1
INV T 15	C4	3	0	0	1	0	0	x	x
INV T 15	C5	8	0	3	1	0	0	5	2
INV T 15	C6	4	0	0	3	0	0	20	7
INV T 15	C7	6	0	0	3	0	0	5	2

INV T 15	D1	8	8	3	0	0	0	8	3
INV T 15	D2	8	7	0	0	0	0	5	2
INV T 15	D3	7	8	2	1	0	0	4	1
INV T 15	D4	8	6	0	1	0	0	3	1
INV T 15	D6	8	6	1	1	0	0	4	1
INV T 15	D7	8	0	3	0	0	0	4	1
INV T 15	E1	8	7	0	0	0	0	0	0
INV T 15	E2	8	8	0	0	0	0	4	1
INV T 15	E3	7	0	3	0	0	0	5	x
INV T 15	E4	6	0	0	1	0	80	30	37
TMA block number	INVASIVE TUMOUR location	ER (Allred)	PR (Allred)	HER2 (HER2)	EGFR (HER2)	CK5/6 (%)	CK14 (%)	KI67 (%)	MCM
INV T 15	E5	8	8	0	1	0	0	5	2
INV T 15	E6	4	0	0	2	0	0	35	12
INV T 15	E7	0	0	0	2	0	0	20	7
INV T 15	F1	3	0	3	0	0	0	x	x
INV T 15	F2	0	3	0	2	5	4	20	10
INV T 15	F3	8	8	3	0	0	0	8	3
INV T 15	F4	8	8	0	0	0	0	5	2
INV T 15	F5	8	4	0	1	0	0	4	1
INV T 15	F7	8	0	0	0	0	0	5	2
INV T 15	G1	4	0	3	1	0	0	15	5
INV T 15	G2	8	0	3	0	0	0	x	x

INV T 15	G3	3	0	0	2	20	0	25	15
INV T 15	G4	0	0	0	2	10	0	40	17
INV T 15	G6	8	0	2	0	0	0	5	2
INV T 15	G7	4	0	0	3	0	0	12	4

5.17 Appendix 3: Tissue Microarray Patient Demographics

Research id number	Surgery type	Side	Recurrence surgery type	Side	Time between specimens (days)	Size (mm)	Margins	Grade	Main architecture	Necrosis	Chronic inflammation	Recurrence y/n	Recurrence type	Size recoded at median of 15mm
100448	Excision biopsy	Left	Radical mastectomy	Left	5	7.0	Not available	High grade	Solid	Marked	Marked	Y	DCIS	Less than 15mm
100453	Excision biopsy	Left	Radical mastectomy	Left	5	16.0	Not available	High grade	Micropapillary	Moderate	Mild	Y	DCIS	15mm or more
100465	Excision biopsy	Right	Simple mastectomy	Right	0	13.0	Not available	High grade	Solid	Mild	Moderate	Y	DCIS	Less than 15mm
100466	Excision biopsy	Right	Simple mastectomy	Right	7	5.0	Not available	High grade	Cribriform	None	None	Y	DCIS	Less than 15mm
100491	Micdo	Right	Simple mastectomy	Right	14	16.0	Not available	High grade	Micropapillary	Moderate	Moderate	Y	DCIS	15mm or more
100494	Excision biopsy	Right	Radical mastectomy	Right	7	29.0	Not available	High grade	Solid	Marked	Marked	Y	DCIS	15mm or more
100508	Excision biopsy	Left	Simple mastectomy	Left	35			No DCIS				Y	DCIS	
100523	Excision biopsy	Right	Wide local excision	Right	62	21.0	Not available	High grade	Cribriform	Marked	None	Y	DCIS	15mm or more
100528	Micdo	Left	Simple mastectomy	Left	24	2.0	Not available	High grade	Papillary	None	None	Y	DCIS	Less than 15mm
100534	Excision biopsy	Left	Simple mastectomy	Left	0	15.0	Not available	High grade	Solid	Marked	Marked	Y	DCIS	15mm or more
100538	Biopsy	Left	Simple mastectomy	Left	9	21.0	Not available	Intermediate grade	Micropapillary	Moderate	Moderate	Y	DCIS	15mm or more
100541	Wide excision	Right	Simple mastectomy	Right	14	20.0	0.0	High grade	Solid	Marked	Moderate	Y	DCIS	15mm or more
100543	Excision biopsy	Right	Simple mastectomy	Right	19	25.0	0.0	High grade	Solid	Marked	Marked	Y	DCIS	15mm or more
100546	Biopsy	Right			0	19.0	0.0	High grade	Solid	Moderate	Mild	N	No	15mm or more
100547	Excision biopsy	Left	Radical mastectomy	Right	0	Na	Not available	Intermediate grade	Cribriform	None	Moderate	Y	DCIS	
100548	Excision biopsy	Right	Wide local excision	Right	21	31.0	0.0	Intermediate grade	Cribriform		Mild	Y	DCIS	15mm or more
Research id number	Surgery type	Side	Recurrence surgery type	Side	Time between specimens	Size	Margins	Grade	Main architecture	Necrosis	Ci	Recurrence y/n	DCIS	Less than 15mm
100555	Excision biopsy	Left	Wide local excision	Left	12	12.0	0.0	Intermediate grade	Cribriform	Mild	Mild	Y	DCIS	Less than 15mm
100561	Excision biopsy	Left	Simple mastectomy	Left	12	18.0	0.0	High grade	Solid	Marked	Marked	Y	DCIS	15mm or more
100562	Excision biopsy	Left	Excision biopsy	Left	2126							Y	DCIS	
100572	Excision biopsy	Left	Radical mastectomy	Left	29	26.0	0.0	High grade	Solid	Marked	Moderate	Y	DCIS	15mm or more
100580	Excision biopsy	Right	Wide local excision	Right	15	16.0	Not available	Intermediate grade	Cribriform	Moderate	None	Y	DCIS	15mm or more
100590	Excision biopsy	Right	Simple mastectomy	Right	26	21.0	Not available	Intermediate grade	Cribriform	Mild	None	Y	DCIS	15mm or more
100591	Biopsy	Left	Radical mastectomy	Left	19	25.0	0.0	Intermediate grade	Cribriform	Marked	Moderate	Y	DCIS	15mm or more

100597	Excision biopsy	Left	Wide local excision	Left	41	22.0	0	High grade	Micropapillary	Marked	Mild	Y	DCIS	Less than 15mm
100601	Excision biopsy	Right	Radical mastectomy	Right	40	8.0	0.0	High grade	Solid	Moderate	Marked	Y	DCIS	Less than 15mm
100602	Referral	Left	Excision biopsy, simple mastectomy	Left	46	3.0	1.0	High grade	Solid	Moderate	Mild	N	N	Less than 15mm
100607	Excision biopsy	Left	Radical mastectomy	Left	27	22.0	0.0	High grade	Cribriform	Moderate	None	Y	DCIS	15mm or more
100609	Excision biopsy	Left	Wide local excision	Left	33	12.0	0.0	High grade	Solid	Marked	Marked	Y	DCIS	Less than 15mm
100611	Excision biopsy	Left	Wide local excision	Left	34	18.0	0.0	Intermediate grade	Cribriform	Mild	Moderate	Y	DCIS	15mm or more
100613	Referral	Right	Excision biopsy	Right	0	11.0	0.0	High grade	Solid	Moderate	Mild	N	N	Less than 15mm
100615	Biopsy	Left	Excision biopsy	Left	33	16.0	0.0	High grade	Cribriform	Moderate	Mild	N	N	15mm or more
100617	Referral	Left	Simple mastectomy	Left	22	21.0	0.0	Intermediate grade	Cribriform	None	None	Y	DCIS	15mm or more
100618	Excision biopsy	Right	Wide local excision	Right	30	17.0	0.0	Intermediate grade	Cribriform	Moderate	Mild	Y	DCIS	15mm or more
100619	Referral	Left	Simple mastectomy	Left	33	21.0	Not available	High grade	Solid	Marked	Marked	N	N	15mm or more
100621	Excision biopsy	Left	Wide local excision	Left	21	10.0	Not available	Intermediate grade	Cribriform	None	Mild	Y	DCIS	Less than 15mm
100630	Referral	Left	Excision biopsy	Left	0	14.0	1.0	Low grade	Cribriform	None	Mild	N	N	Less than 15mm
Research id number	Surgery type	Side	Recurrence surgery type	Side	Time between specimens (days)	Size (mm)	Margins	Grade	Main architecture	Necrosis	Chronic inflammation	Recurrence y/n	Recurrence type	Size recoded at median of 15mm
100631	Referral		Simple mastectomy	Left	769	15.0	Not available	High grade	Cribriform	Moderate	Moderate	N	N	15mm or more
100633	Excision biopsy	Right	Wide local excision	Right	36	6.0	0.0	Intermediate grade	Cribriform	None	Na	No rid	DCIS	Less than 15mm
100639	Excision biopsy	Right	Wide local excision	Right	27	15.0	0.0	High grade	Solid	Marked	Marked	No research id number	DCIS	15mm or more

100643	Referral		Excision biopsy	Left	226	20.0	0.0	Intermediate grade	Cribriform	Moderate	Moderate	N	N	15mm or more
100644	Excision biopsy	Right	Wide local excision	Right	43	10.0	0.0	High grade	Cribriform	Mild	Moderate	Y	DCIS	Less than 15mm
100645	Micdo	Right	Simple mastectomy	Right	29	15.0	0.0	High grade	Micropapillary	Mild	Moderate	No research id number		15mm or more
100646	Referral	Left	Excision biopsy	Left	108							Y	DCIS	
100649	Referral	Right	Excision biopsy	Right	36	10.0	Not available	Intermediate grade	Cribriform	Mild	Mild	N	N	Less than 15mm
100650	Referral	Right	Excision biopsy	Right	36	5.0	1.5	High grade	Solid	Mild	Marked	N	N	Less than 15mm
100651	Excision biopsy	Left	Wide local excision	Left	126	16.0	0.0	High grade	Solid	Marked	Marked	No research id number		15mm or more
100658	Excision biopsy	Left	Wide local excision, simple mastectomy	Left	63	17.0	Not available	High grade	Cribriform	Marked	Moderate	Y	DCIS	15mm or more
100662	Excision biopsy	Right	Simple mastectomy	Right	54	14.0	0.0	High grade	Cribriform	Marked	Marked	Y	DCIS	Less than 15mm
100663	Referral	Right	Simple mastectomy	Right	80	7.0	Not available	Low grade	Cribriform	None	None	N	N	Less than 15mm
100667	Micdo	Right	Simple mastectomy	Right	49	23.0	0.0	High grade	Solid	Marked	Mild	Y	DCIS	15mm or more
100671	Excision biopsy	Left	Wide local excision, simple mastectomy	Left	34	14.0	0.0	Intermediate grade	Cribriform	Mild	Moderate	Y	DCIS	Less than 15mm
100674	Excision biopsy	Right	Wide local excision, simple mastectomy	Right left	50	9.0	0.0	High grade	Cribriform	Mild	Moderate	Y	DCIS	Less than 15mm

NOTE: All patient identifiers and dates have been removed.

Table 32: Patient Demographics: RID= Research Number

Research id number	Surgery type	Side	Recurrence surgery type	Side	Time between specimens (days)	Size (mm)	Margins	Grade	Main architecture	Necrosis	Chronic inflammation	Recurrence y/n	Recurrence type	Size recoded at median of 15mm
100676	Excision biopsy	Right	Wide local excision. Simple mastectomy	Right	36	15.0	Not available	High grade	Micropapillary	Moderate	Marked	Y	DCIS	15mm or more
100679	Excision biopsy	Right	Simple mastectomy	Right	43	27.0	0.0	High grade	Micropapillary	Marked	Marked	Y	DCIS	15mm or more
100696	Micdo	Left	Simple mastectomy	Left	55	8.0	Not available	High grade	Cribriform	None	Mild	N	N	Less than 15mm
100699	Excision biopsy	Left	Simple mastectomy	Left	63			No DCIS						
100701	Excision biopsy	Right	Simple mastectomy	Right left	34	1.0	Not available	High grade	Solid	None	None	N	N	Less than 15mm
100703	Excision biopsy	Right	Simple mastectomy	Right	28	20.0	Not available	High grade	Cribriform	Mild	Mild	No research id number		15mm or more
													N	
117057	Micdo	Right	Simple mastectomy	Right	27	19.0	0.0	High grade	Solid	None	Mild	Y	DCIS	15mm or more
117066			Wide local excision	Right	27	22.0	0.0	Intermediate grade	Cribriform	Mild	Mild	N	N	15mm or more
117079	Excision biopsy	Left	Simple mastectomy	Left	40	16.0	Not available	Intermediate grade	Micropapillary	None	Mild	Y	DCIS	15mm or more
117081	Excision biopsy	Left	Wide local excision	Left	21	20.0	Not available	Intermediate grade	Cribriform	Mild	Mild	Y	DCIS	15mm or more
117088	Excision biopsy	Right	Wide local excision	Right	21	24.0	0.0	High grade	Solid	Marked	Marked	Y	DCIS	15mm or more
117092	Excision biopsy	Left	Simple mastectomy	Left	22	18.0	Not available	Intermediate grade	Papillary	None	Mild	N	N	15mm or more
117098	Micdo	Left	Wide local excision & ancillary clearance	Left	61	Na	Not available	Intermediate grade	Papillary	None	None	Y	DCIS	
117105	Referral	Right	Excision biopsy	Right	?	11.0	Not available	Intermediate grade	Solid	Marked	None	N	N	Less than 15mm
117119	Excision biopsy	Right	Wide local excision	Right	21	12.0	0.0	Intermediate grade	Micropapillary	None	None	Y	DCIS	Less than 15mm
117122	Referral	Right	Simple mastectomy	Right	?	Na	Not available	Low grade	Cribriform	None	None	N	N	
117124	Referral	Left	Simple mastectomy	Left	14	29.0	Not available	High grade	Cribriform	Mod	Mild	N	N	15mm or more

Research id number	Surgery type	Side	Recurrence surgery type	Side	Time between specimens (days)	Size (mm)	Margins	Grade	Main architecture	Necrosis	Chronic inflammation	Recurrence y/n	Recurrence type	Size recoded at median of 15mm
117125	Excision biopsy	Right	Wide local excision	Right	58	13.0	0.0	Intermediate grade	Cribriform	None	Moderate	No id number	DCIS	Less than 15mm
117130	Referral		Excision biopsy& wide local excision	Left	?	6.0	3.0	High grade	Solid	Marked	Marked	N	N	Less than 15mm
117145	Referral	Right	Simple mastectomy	Right	?	16.0	Not available	High grade	Cribriform	Marked	Moderate	N	N	15mm or more
117147	Wle	Right	Simple mastectomy	Right	62	19.0	Not available	High grade	Solid	Marked	Marked	Y	DCIS	15mm or more
117148	Excision biopsy	Left	Referral	Left	0	9.0	Not available	Intermediate grade	Papillary	None	Mild	Y	DCIS	Less than 15mm
117152	Referral	Left	Simple mastectomy	Left	?	24.0	Not available	High grade	Solid	Mod	None	N	N	15mm or more
117155	Excision biopsy	Right	Wide local excision	Right	20	13.0	0.0	Intermediate grade	Solid	Mild	None	Y	DCIS	Less than 15mm
117159	Wle	Right	Radical mastectomy	Right	26	17.0	Not available	High grade	Solid	Mild	Marked	Y	DCIS	15mm or more
117163	Excision biopsy	Right	Simple mastectomy	Right	28	16.0	0.0	Intermediate grade	Cribriform	Mod	None	Y	DCIS	15mm or more
117174	Sm	Left	Referral	Left	0	21.0	Not available	High grade	Solid	Mod	Marked	Y	DCIS	15mm or more
117176	Referral	Left	Excision biopsy	Left	?	21.0	Not available	Intermediate grade	Solid	Mild	Mild	N	N	
117178	Excision biopsy	Left	Simple mastectomy	Left	14	19.0	0.0	Low grade	Cribriform	None	Mild	Y	DCIS	15mm or more
117181	Excision biopsy	Left	Referral & simple mastectomy	Left	14	20.0	0.0	High grade	Solid	Mod	Moderate	Y	DCIS	15mm or more
117183	Referral	Left	Excision biopsy	Left	?	8.0	Not available	High grade	Solid	None	Mild	N	N	Less than 15mm
117184	Excision biopsy	Right	Ref & wide local excision	Right	43	Na	Not available	High grade	Cribriform	None	Moderate	Y	DCIS	
117185	Referral	Left	Excision biopsy & simple mastectomy	Left	?	27.0	0.0	Intermediate grade	Cribriform	Mild	Mild	N	N	15mm or more
117195	Excision biopsy	Left	Wide local excision	Left	15	19.0	0.0	High grade	Solid	Mod	Marked	Y	DCIS	15mm or more
117204	Excision biopsy	Right	Wide local excision	Right	23	8.0	1.0	Low grade	Cribriform	None	None	No research id number		Less than 15mm
117208	Excision biopsy	Right	Wide local excision	Right	23	27.0	0.0	Intermediate grade	Papillary	Mild	Moderate	Y	DCIS	15mm or more

Table 33: *IHC* scores for TMA cases with associated patient data.

Research id number	Final HER2 status	Final egfr positive score (any)	Final average er score	Final er status	Final average pr scores	Final pr status	Final highest ck5 scores	Final positive ck5 s (more than 1%)	Final highest ck5/6 score	Ck 5/6 positive (more than 1%)	Final highest ck14 score	Ck14 positive (more than 1%)	Ki67 tma average	Ki67 positive more than 5% (median)	MCM2 tma average (%)	MCM2 Positive (median 27.5%)
100448	0	1	0	0	0	0	0.5	0	1.0	1	1.0	1	7	1	33	1
100453	0	0	5	1	1	0	0.0	0	5.0	1	0.5	0	13	1	40	1
100465	1	1	1	0	0	0	0.5	0	1.0	1	0.5	0	13	1	28	1
100466																
100491	1	0	0	0	0	0	0.0	0	0.0	0	0.0	0	17	1	40	1
100494	1	0	1	0	0	0	1.0	0	5.0	1	5.0	1	45	1	91	1
100508	0	0			2	0			1.0	1	5.0	1	40	1	35	1
100523	0	0	7	1	2	0	0.0	0	0.0	0	0.0	0	10	1	5	0
100528																
100534	1	0	0	0	0	0	0.5	0	2.0	1	20.0	1	13	1	88	1
100538	0	0	4	1	0	0	0.0	0	0.0	0	0.5	0	8	1	30	1
100541	1	0	3	0	0	0	2.0	1	0.0	0	5.0	1	28	1	85	1
100543	1	0	1	0	0	0	0.5	0	1.0	1	2.0	1	18	1	72	1
100546	0	0	8	1	6	1	0.0	0	0.0	0	1.0	1	0	0	4	0
100547	0	0	8	1	6	1	0.0	0	1.0	1	0.0	0	1	0	25	0
100548	1	0	8	1	0	0	0.5	0	0.0	0	0.5	0	2	0	13	0
100550	0	0	6	1	2	0	0.0	0	1.0	1	10.0	1	6	1	48	1
100555	0	0	8	1	7	1	0.5	0	0.0	0	1.0	1	5	1	25	0
100561	1	1	0	0	0	0	0.5	0	0.0	0	0.0	0	7	1	78	1
100562	0				0	0	0.0	0			0.0	0	5	1	5	0

100572	1	0	8	1	0	0	0.5	0	1.0	1	1.0	1	30	1	50	1
Research id number	Final HER2 status	Final egfr positive score (any)	Final average er score	Final er status	Final average pr scores	Final pr status	Final highest ck5 scores	Final positive ck5 s (more than 1%)	Final highest ck5/6 score	Ck 5/6 positive (more than 1%)	Final highest ck14 score	Ck14 positive (more than 1%)	Ki67 tma average	Ki67 positive more than 5% (median)	MCM2 tma average (%)	MCM2 Positive (median 27.5%)
100580	0	0	7	1	7	1	0.0	0	0.0	0	0.0	0	5	1	15	0
100590	0	0	7	1	2	0	0.0	0	0.0	0	0.0	0	6	1	0	0
100597	0	0	8	1	6	1	0.0	0	0.0	0	0.0	0	5	1	13	0
100601	1				0	0	0.0	0			0.0	0			0	0
100602	0	0	8	1	4	1	0.0	0	0.0	0	0.0	0	10	1	70	1
100606	0	0	6	1	6	1	1.0	0	0.5	0	0.5	0	5	1	23	0
100607	1	0	0	0	0	0	1.0	0	10.0	1	5.0	1	40	1	30	1
100609	0	1	0	0	0	0	60.0	1	50.0	1	0.0	0	15	1	18	0
100611	1	0	3	1	0	0	1.0	0	0.0	0	0.0	0	0	0	30	
100613	1	0	5	1	0	0	0.0	0			0.0	0	1	0	35	1
100615	0	0	4	1	0	0	0.0	0	0.0	0	0.0	0	1	0	9	0
100617	0	0	8	1	2	0	0.0	0	0.0	0	0.0	0	0	0	16	0
100618	0	0	8	1	8	1	0.0	0	1.0	1	0.0	0	2	0	15	0
100619	1	0	0	0	0	0	1.0	0	1.0	1	1.0	1	25	1	80	1
100621	0	0	7	1	8	1	1.0	0	0.0	0	0.0	0	7	1	20	0
100630	0	0	7	1	3	1	0.0	0	0.0	0	0.0	0	0	0	5	0
100631	0	1	5	1	1	0	5.0	1	0.0	0	0.0	0	0	0	63	1
100633	0	0	8	1	4	1	0.0	0	0.0	0	0.0	0	0	0	8	0
100637	0	0	6	1	0	0	0.0	0	2.0	1	0.0	0	4	0	53	1
100639	1	0	7	1	5	1	2.0	1	1.0	1	0.0	0	8	1	88	1

100643	0	0	8	1	7	1	0.0	0	0.0	0	0.0	0	0	0	43	1
100644	0	0	8	1	7	1	1.0	0	0.0	0	0.0	0	3	0	30	1
100645	1	0	0	0	0	0	0.0	0	0.0	0	0.0	0	5	1	27	0
100646	1	0	0	0	0	0	1.0	0	0.0	0	0.0	0	10	1	45	1

Research id number	Final HER2 status	Final egfr positive score (any)	Final average er score	Final er status	Final average pr scores	Final pr status	Final highest ck5 scores	Final positive ck5 s (more than 1%)	Final highest ck5/6 score	Ck 5/6 positive (more than 1%)	Final highest ck14 score	Ck14 positive (more than 1%)	Ki67 tma average	Ki67 positive more than 5% (median)	MCM2 tma average (%)	MCM2 Positive (median 27.5%)
100649	0	0	8	1	8	1	0.0	0	0.0	0	0.0	0	5	1	65	1
100650	1	0	0	0	0	0	0.0	0	0.0	0	0.0	0	10	1	70	1
100651	1	0	0	0	0	0	10.0	1	15.0	1	5.0	1	10	1	50	1
100662	1	0	0	0	0	0	0.0	0	1.0	1	0.0	0	7	1	47	1
100663	0	0	8	1	6	1	0.0	0	5.0	1	0.0	0	0	0	7	0
100667	0	0	6	1	6	1	0.0	0	0.0	0	0.0	0	18	1	32	1
100671	1	0	8	1	7	1	2.0	1	1.0	1	0.0	0	10	1	32	1
100674	0	0	6	1	3	1	2.0	1	10.0	1	5.0	1	8	1	23	0
100676	0	0	0	0	0	0	1.0	0	0.0	0	0.0	0	6	1	23	0
100679	1	0	8	1	1	0	1.0	0	1.0	1	0.0	0	5	1	42	1
100696	0	0			7	1	0.0	0	0.0	0	0.0	0	3	0	13	0
100699																
100701																
100703	0	0	8	1	7	1	0.0	0	0.0	0	0.0	0	3	0	29	1
117044	1	0	0	0	0	0	1.0	0	1.0	1	1.0	1	10	1	13	0
117057	0	0	8	1	8	1	5.0	1	2.0	1	2.0	1	0	0	20	0
117061	0	0	8	1	5	1	0.0	0	0.0	0	0.0	0	0	0	10	0
117066	0	0	6	1	0	0	0.0	0	0.0	0	0.0	0	8	1	10	0
117079	0	0	0	0	2	0	5.0	1	1.0	1	0.0	0	10	1	0	0
117081	0	0	7	1	0	0	0.0	0	1.0	1	0.0	0	5	1	27	0
117088	1	0	0	0	0	0	1.0	0	2.0	1	1.0	1	23	1	65	1

117092	0	0	8	1	7	1	1.0	0	0.5	0	0.5	0	3	0	11	0
117098	0	0	8	1	2	0	0.0	0	0.0	0	0.0	0	0	0	0	0

Research id number	Final HER2 status	Final egfr positive score (any)	Final average er score	Final er status	Final average pr scores	Final pr status	Final highest ck5 scores	Final positive ck5 s (more than 1%)	Final highest ck5/6 score	Ck 5/6 positive (more than 1%)	Final highest ck14 score	Ck14 positive (more than 1%)	Ki67 tma average	Ki67 positive more than 5% (median)	MCM2 tma average (%)	MCM2 Positive (median 27.5%)
117105	0	0	8	1	4	1	0.0	0	0.0	0	0.0	0	0	0	10	0
117119	0	0			0	0	0.0	0			0.0	0	1	0	1	0
117122	0				8	1							0	0	1	0
117124	0	0	8	1	4	1	1.0	0	0.0	0	0.0	0	2	0	40	1
117125	0	0	8	1		1	2.0	1	0.0	0	0.0	0	0	0	10	0
117145	0	0	8	1	6	1	0.0	0	0.0	0	0.0	0	2	0	28	1
117147	1	1	1	0	0	0	0.0	0	0.0	0	0.0	0	12	1	82	1
117148	0	0	8	1	8	1	0.0	0	0.0	0	0.0	0				
117152	0	0	8	1	4	1	0.0	0	0.0	0	0.0	0	2	0	10	0
117155	0	0			5	1			1.0	1	0.0	0	2	0	25	0
117159	0	0	0	0	0	0	0.0	0	0.0	0	0.0	0	3	0	35	1
117163	0	0	8	1	6	1	0.0	0	0.0	0	0.0	0	5	1	57	1
117164	0	0	8	1	1	0	5.0	1	5.0	1	5.0	1	20	1	87	1
117174	1	0	0	0	0	0	1.0	0	1.0	1	0.0	0	27	1	75	1
117176	0		8	1	8	1	0.0	0	0.0	0	0.0	0				
117178	0	0	8	1	8	1	0.0	0	1.0	1	0.0	0	5	1	43	1
117181	0	0	8	1	4	1	5.0	1	3.0	1	5.0	1	7	1	23	0
117183	0	0	5	1	5	1	0.0	0	0.0	0	0.0	0			40	1
117184	0	0	6	1	1	0	0.0	0	0.0	0	0.0	0	0	0	10	0
117185	0	0			6	1							0	0	10	0
117195	0	0	8	1	5	1	0.0	0	0.0	0	0.0	0	8	1	35	1
117204	0	0	8	1	7	1	0.0	0	0.5	0	0.0	0	1	0	29	1

117208	0	0	8	1	0	0	0.0	0	0.0	0	0.0	0	4	0	33	1
--------	---	---	---	---	---	---	-----	---	-----	---	-----	---	---	---	----	---

5.18 Appendix 4: Immunohistochemistry Protocols

- i) Dewax slides in dewax solution followed by alcohol.
- ii) Epitope retrieval solution according to antibody.
- iii) Peroxidase block, 5 mins.
- iv) Primary antibody incubation, 15 mins.
- v) Post-primary, 8 mins.
- vi) Polymer, 8 mins.
- vii) DAB, 10 mins.
- viii) DAB Enhancer, 10 mins.
- ix) Haematoxylin counterstain, 5 mins.
- x) Rinse in water, dehydrate, clear and mount manually.

Bond wash buffer rinses were carried out by the machine between each of these steps.

Antibodies were diluted using Bond Antibody Diluent. Bond epitope retrieval solutions 1, & 2 were used where stated at a temperature of 95°C

Antibody	Dilution	Epitope retrieval soln	Epitope retrieval incubation (minutes)
ER 6F11	1/100	ER1	30
PR	1/400	ER1	30
CK5/6	1/100	ER1	30
CK5	1/100	ER1	30
CK14	1/100	ER1	30
EGFR	1/50	ER1	20

HER2 IHC was performed using the Bond “Oracle” HER2 kit as per manufacturers instructions.

5.19 Appendix 5: Frequency Tables and Non Parametric Statistical Analysis of DCIS IHC

Frequency Distribution for FINAL HER2 STATUS
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
0	191	77.016
1	57	22.984
Total	248	100.000

Frequency Distribution for FINAL EGFR STATUS
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
0	216	95.154
1	11	4.846
Total	227	100.000

Frequency Distribution for FINAL ER STATUS
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
0	59	26.339
1	165	73.661
Total	224	100.000

Frequency Distribution for FINAL PR STATUS
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
0	123	52.119
1	113	47.881
Total	236	100.000

Frequency Distribution for FINAL CK5 (AT 1%)
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
0	195	85.903
1	32	14.097
Total	227	100.000

Frequency Distribution for FINAL CK5/6 (AT 1%)
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
0	151	68.636
1	69	31.364
Total	220	100.000

Frequency Distribution for FINAL CK14 (AT 1%)
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
0	193	83.550
1	38	16.450
Total	231	100.000

Frequency Distribution for GRADE
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
HG	98	57.988
IG	58	34.320
LG	13	7.692
Total	169	100.000

Frequency Distribution for ARCH
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
CRIB	68	40.237
MICROPAP	18	10.651
PAP	13	7.692
SOLID	70	41.420
Total	169	100.000

Frequency Distribution for NECROSIS
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
MARKED	42	24.852
MILD	47	27.811
MOD	38	22.485
NONE	42	24.852
Total	169	100.000

Frequency Distribution for CI
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
MARKED	37	22.156
MILD	59	35.329
MOD	36	21.557
NONE	35	20.958
Total	167	100.000

Frequency Distribution for REC Y/N
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
N	140	85.366
Y	24	14.634
Total	164	100.000

Frequency Distribution for REC TYPE
Row exclusion: DCIS STATVIEW DATASET

	Count	Percent
BILATERAL INVASIVE CARCINOMA	1	.613
CONTRALATERAL INVASIVE CARCINOMA	3	1.840
IPSILATERAL DCIS	6	3.681
IPSILATERAL INVASIVE CA	7	4.294
NK	6	3.681
NONE	140	85.890
Total	163	100.000

Summary Table for FINAL K167 (AT MEDIAN, 5%), FINAL HER2 STATUS

Row exclusion: DGIS STATVIEW DATASET

Num. Missing	43
DF	1
Chi Square	15.451
Chi Square P-Value	<.0001
G-Squared	16.560
G-Squared P-Value	<.0001
Contingency Coef.	.250
Phi	.258
Cty. Cor. Chi Square	14.255
Cty. Cor. P-Value	.0002
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL K167 (AT MEDIAN, 5%), FINAL HER2 STATUS

Row exclusion: DGIS STATVIEW DATASET

	0	1	Totals
0	87	11	98
1	89	45	134
Totals	176	56	232

Summary Table for FINAL EGFR STATUS, FINAL K167 (AT MEDIAN, 5%)

Row exclusion: DGIS STATVIEW DATASET

Num. Missing	53
DF	1
Chi Square	4.869
Chi Square P-Value	.0273
G-Squared	5.872
G-Squared P-Value	.0154
Contingency Coef.	.146
Phi	.148
Cty. Cor. Chi Square	3.580
Cty. Cor. P-Value	.0585
Fisher's Exact P-Value	.0294

Observed Frequencies for FINAL EGFR STATUS, FINAL K167 (AT MEDIAN, 5%)

Row exclusion: DGIS STATVIEW DATASET

	0	1	Totals
0	90	121	211
1	1	10	11
Totals	91	131	222

Summary Table for FINAL ER STATUS, FINAL K67 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	59
DF	1
Chi Square	23.351
Chi Square P-Value	<.0001
G-Square	25.940
G-Square P-Value	<.0001
Contingency Coef.	.312
Phi	.329
Cty. Cor. Chi Square	21.868
Cty. Cor. P-Value	<.0001
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL ER STATUS, FINAL K67 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	8	51	59
1	78	79	157
Totals	86	130	216

Summary Table for FINAL PR STATUS, FINAL K67 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	51
DF	1
Chi Square	15.527
Chi Square P-Value	<.0001
G-Square	15.684
G-Square P-Value	<.0001
Contingency Coef.	.255
Phi	.263
Cty. Cor. Chi Square	14.479
Cty. Cor. P-Value	.0001
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL PR STATUS, FINAL K67 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	36	82	118
1	60	46	106
Totals	96	128	224

Summary Table for FINAL CKS (AT 1%), FINAL K67 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	58
DF	1
Chi Square	1.222
Chi Square P-Value	.2690
G-Square	1.254
G-Square P-Value	.2629
Contingency Coef.	.075
Phi	.075
Cty. Cor. Chi Square	.830
Cty. Cor. P-Value	.3623
Fisher's Exact P-Value	.3305

Observed Frequencies for FINAL CKS (AT 1%), FINAL K67 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	77	108	185
1	10	22	32
Totals	87	130	217

Summary Table for FINAL CK5/6 (AT 1%), FINAL K067 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	63
DF	1
Chi Square	1.733
Chi Square P-Value	.1881
G-Squared	1.756
G-Squared P-Value	.1851
Contingency Coef.	.090
Phi	.090
Cty. Cor. Chi Square	1.358
Cty. Cor. P-Value	.2439
Fisher's Exact P-Value	.2279

Observed Frequencies for FINAL CK5/6 (AT 1%), FINAL K067 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	59	84	143
1	22	47	69
Totals	81	131	212

Summary Table for FINAL CK14 (AT 1%), FINAL K067 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	54
DF	1
Chi Square	4.985
Chi Square P-Value	.0256
G-Squared	5.275
G-Squared P-Value	.0216
Contingency Coef.	.149
Phi	.150
Cty. Cor. Chi Square	4.209
Cty. Cor. P-Value	.0402
Fisher's Exact P-Value	.0290

Observed Frequencies for FINAL CK14 (AT 1%), FINAL K067 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	79	104	183
1	9	29	38
Totals	88	133	221

Summary Table for FINAL MCM2 (AT MEDIAN, 27.5%), FINAL K067 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	38
DF	1
Chi Square	41.483
Chi Square P-Value	<.0001
G-Squared	42.558
G-Squared P-Value	<.0001
Contingency Coef.	.386
Phi	.418
Cty. Cor. Chi Square	39.807
Cty. Cor. P-Value	<.0001
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL MCM2 (AT MEDIAN, 27.5%), FINAL K067 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	70	37	107
1	31	99	130
Totals	101	136	237

SummaryTable for GRADE, FINAL K167 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Mssing	121
DF	2
Chi Square	13.981
Chi Square P-Value	.0009
G-Squared	14.052
G-Squared P-Value	.0009
Contingency Coef.	.288
Cramer's V	.301

Observed Frequencies for GRADE, FINAL K167 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
HG	26	65	91
IG	28	24	52
LG	8	3	11
Totals	62	92	154

SummaryTable for ARCH, FINAL K167 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Mssing	121
DF	3
Chi Square	12.020
Chi Square P-Value	.0073
G-Squared	12.697
G-Squared P-Value	.0053
Contingency Coef.	.269
Cramer's V	.279

Observed Frequencies for ARCH, FINAL K167 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
CRIB	34	30	64
MICROPAP	2	13	15
PAP	6	5	11
SOLID	20	44	64
Totals	62	92	154

SummaryTable for NECROSIS, FINAL K167 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Mssing	121
DF	3
Chi Square	27.774
Chi Square P-Value	<.0001
G-Squared	28.911
G-Squared P-Value	<.0001
Contingency Coef.	.391
Cramer's V	.425

Observed Frequencies for NECROSIS, FINAL K167 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
MARKED	7	35	42
MILD	18	25	43
MOD	12	24	36
NONE	25	8	33
Totals	62	92	154

SummaryTable for CI, FINAL K167 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Mssing	123
DF	3
Chi Square	21.210
Chi Square P-Value	<.0001
G-Squared	24.603
G-Squared P-Value	<.0001
Contingency Coef.	.350
Cramer's V	.374

Observed Frequencies for CI, FINAL K167 (AT MEDIAN, 5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
MARKED	3	33	36
MILD	27	26	53
MOD	15	20	35
NONE	16	12	28
Totals	61	91	152

SummaryTable for FINAL MCM2 (AT MEDIAN, 27.5%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

Num. Mssing	39
DF	1
Chi Square	19.314
Chi Square P-Value	<.0001
G-Squared	20.682
G-Squared P-Value	<.0001
Contingency Coef.	.275
Phi	.286
Cy. Cor. Chi Square	17.992
Cy. Cor. P-Value	<.0001
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL MCM2 (AT MEDIAN, 27.5%), FINAL HER2 STATUS

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	94	11	105
1	85	46	131
Totals	179	57	236

SummaryTable for FINAL EGFR STATUS, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Mssing	50
DF	1
Chi Square	5.588
Chi Square P-Value	.0181
G-Squared	6.671
G-Squared P-Value	.0098
Contingency Coef.	.156
Phi	.158
Cy. Cor. Chi Square	4.217
Cy. Cor. P-Value	.0400
Fisher's Exact P-Value	.0254

Observed Frequencies for FINAL EGFR STATUS, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	97	117	214
1	1	10	11
Totals	98	127	225

Summary Table for FINAL ER STATUS, RNAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET		
Num. Missing		56
DF		1
Chi Square		9.119
Chi Square P-Value		.0025
G-Squared		9.500
G-Squared P-Value		.0021
Contingency Coef.		.200
Phi		.204
Cty. Cor. Chi Square		8.215
Cty. Cor. P-Value		.0042
Fisher's Exact P-Value		.0032

Observed Frequencies for FINAL ER STATUS, RNAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET			
	0	1	Totals
0	15	44	59
1	77	83	160
Totals	92	127	219

Summary Table for FINAL PR STATUS, RNAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET		
Num. Missing		47
DF		1
Chi Square		8.977
Chi Square P-Value		.0027
G-Squared		9.029
G-Squared P-Value		.0027
Contingency Coef.		.195
Phi		.198
Cty. Cor. Chi Square		8.186
Cty. Cor. P-Value		.0042
Fisher's Exact P-Value		.0033

Observed Frequencies for FINAL PR STATUS, RNAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET			
	0	1	Totals
0	42	77	119
1	60	49	109
Totals	102	126	228

Summary Table for FINAL CK5 (AT 1%), RNAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET		
Num. Missing		54
DF		1
Chi Square		.912
Chi Square P-Value		.3396
G-Squared		.928
G-Squared P-Value		.3353
Contingency Coef.		.064
Phi		.064
Cty. Cor. Chi Square		.581
Cty. Cor. P-Value		.4460
Fisher's Exact P-Value		.4392

Observed Frequencies for FINAL CK5 (AT 1%), RNAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET			
	0	1	Totals
0	82	107	189
1	11	21	32
Totals	93	128	221

Summary Table for FINAL CK5/6 (AT 1%), RNAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET		
Num. Missing		60
DF		1
Chi Square		.898
Chi Square P-Value		.3434
G-Squared		.904
G-Squared P-Value		.3418
Contingency Coef.		.064
Phi		.065
Cty. Cor. Chi Square		.641
Cty. Cor. P-Value		.4235
Fisher's Exact P-Value		.3772

Observed Frequencies for FINAL CK5/6 (AT 1%), RNAL MCM2 (AT MEDIAN, 27.5%)

Observed Frequencies for FINAL CR5% (At 1%),

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	65	81	146
1	26	43	69
Totals	91	124	215

Summary Table for FINAL CK14 (AT 1%), FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	50
DF	1
Chi Square	1.624
Chi Square P-Value	.2025
G-Squared	1.655
G-Squared P-Value	.1983
Contingency Coef.	.085
Phi	.085
Cty. Cor. Chi Square	1.201
Cty. Cor. P-Value	.2732
Fisher's Exact P-Value	.2151

Observed Frequencies for FINAL CK14 (AT 1%), FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	85	102	187
1	13	25	38
Totals	98	127	225

Summary Table for FINAL MCM2 (AT MEDIAN, 5%), FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	38
DF	1
Chi Square	41.483
Chi Square P-Value	<.0001
G-Square d	42.558
G-Square d P-Value	<.0001
Contingency Coef.	.386
Phi	.418
Cy. Cor. Chi Square	38.807
Cy. Cor. P-Value	<.0001
Fisher's Exact P-Value	<.0001

Observed Frequencies for FINAL MCM2 (AT MEDIAN, 5%), FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
0	70	31	101
1	37	99	136
Totals	107	130	237

Summary Table for GRADE, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	117
DF	2
Chi Square	21.582
Chi Square P-Value	<.0001
G-Square d	21.981
G-Square d P-Value	<.0001
Contingency Coef.	.347
Cramer's V	.370

Observed Frequencies for GRADE, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
HG	28	66	94
IS	36	17	53
LG	7	4	11
Totals	71	87	158

Summary Table for ARCH, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	117
DF	3
Chi Square	11.396
Chi Square P-Value	.0098
G-Square d	11.601
G-Square d P-Value	.0089
Contingency Coef.	.259
Cramer's V	.269

Observed Frequencies for ARCH, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
CRIB	38	27	65
MICROPAP	7	8	15
PAP	6	5	11
SOLID	20	47	67
Totals	71	87	158

Summary Table for NECROSIS, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	117
DF	3
Chi Square	26.520
Chi Square P-Value	<.0001
G-Square d	27.928
G-Square d P-Value	<.0001
Contingency Coef.	.379
Cramer's V	.410

Observed Frequencies for NECROSIS, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
MARKED	9	33	42
MILD	21	24	45
MOD	14	23	37
NONE	27	7	34
Totals	71	87	158

Summary Table for CI, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

Num. Missing	119
DF	3
Chi Square	19.049
Chi Square P-Value	.0003
G-Square d	20.421
G-Square d P-Value	.0001
Contingency Coef.	.330
Cramer's V	.349

Observed Frequencies for CI, FINAL MCM2 (AT MEDIAN, 27.5%)

Row exclusion: DCIS STATVIEW DATASET

	0	1	Totals
MARKED	6	31	37
MILD	31	24	55
MOD	14	21	35
NONE	18	11	29
Totals	69	87	156

Section 2: Non Parametric Tests- Univariate Paired analysis:

Oestrogen receptor in DCIS, frequencies and paired univariate analysis.

ER and HER2 IHC

Fishers exact test p value < 0.0001

	ER -ve	ER +ve	Total
HER2 -ve	23	145	168
HER2 +ve	36	19	55
Total	59	164	223

ER and EGFR IHC

Fishers exact test p value < 0.0001.

	ER -ve	ER +ve	Total
EGFR -ve	45	159	204
EGFR +ve	10	1	11
Total	55	160	215

ER and PR IHC

Fishers exact test p value <0.0001.

	ER -ve	ER +ve	Total
PR -ve	56	55	111
PR +ve	0	102	102
Total	56	157	213

ER and CK5 IHC

Fishers exact test p value = 0.0033.

	ER -ve	ER +ve	Total
CK5 -ve	41	147	188
CK5 +ve	15	16	31
Total	56	163	219

ER and CK5/6 IHC

Fishers exact test p value = 0.298

	ER -ve	ER +ve	Total
CK5/6 -ve	33	114	147
CK5/6 +ve	25	41	66
Total	58	155	213

ER and CK14 IHC

Fishers exact test p value = 0.0018.

	ER -ve	ER +ve	Total
CK14 -ve	40	142	182
CK14 +ve	18	19	37
Total	58	161	219

ER and Ki67 (>5%) IHC

Fishers exact test p value <0.0001

	ER -ve	ER +ve	Total
Ki67 -ve	8	78	86
Ki67+ve	51	79	130
Total	59	157	216

ER and MCM2 (>5%) IHC

Fishers exact test p value = 0.0032

	ER -ve	ER +ve	Total
MCM2 -ve	15	77	92
MCM2+ve	44	83	127
Total	59	160	219

ER and Histological Grade

Chi Square p value <0.0001

	ER -ve	ER +ve	Total
HG	35	55	90
IG	4	46	50
LG	0	9	9
Total	39	110	149

ER and Histological Architecture

Chi square p Value < 0.0001

	ER -ve	ER +ve	Total
Crib	5	55	60
Micropap	7	6	13
Pap	1	11	12
Solid	26	38	64
Total	39	110	149

ER and Necrosis

	ER -ve	ER +ve	Total
Marked	23	18	41
Mild	6	37	43
Moderate	8	28	36
None	2	27	29
Total	39	110	149

ER and Chronic Inflammation

	ER -ve	ER +ve	Total
Marked	25	10	35
Mild	3	49	52
Moderate	9	25	34
None	1	25	26
Total	38	109	147

Progesterone Receptor in DCIS, frequencies and paired univariate analysis.

PR and HER2

Fishers exact test p value <0.0001

	HER2 -ve	HER2 +ve	Total
PR -ve	70	50	120
PR +ve	104	6	110
Total	174	56	230

PR and EGFR IHC

Fishers exact test p value = 0.0008

	PR -ve	PR +ve	Total
EGFR -ve	102	105	207
EGFR +ve	11	0	11
Total	113	105	218

PR and CK5/6 IHC

Fishers exact test p value = 0.0542

	PR -ve	PR +ve	Total
CK5/6 -ve	69	76	145
CK5/6 +ve	42	25	67
Total	111	101	212

PR and CK5 IHC

Fishers exact test p value = 0.2385

	PR -ve	PR +ve	Total
CK5 -ve	94	92	186
CK5 +ve	19	11	30
Total	113	103	216

PR and CK14 IHC

Fishers exact test p value = 0.2791

	PR -ve	PR +ve	Total
CK14 -ve	94	90	184
CK14 +ve	23	14	37
Total	117	104	221

PR and Ki67 (>5%) IHC

Fishers exact test p value <0.0001

	PR -ve	PR +ve	Total
Ki67 -ve	36	60	96
Ki67+ve	82	46	128
Total	118	106	224

PR and MCM2 (<5%) IHC

Chi squared p value = 0.0027

	PR -ve	PR +ve	Total
MCM2 -ve	42	60	102
MCM2+ve	77	49	126
Total	119	109	228

PR and Histological Grade

Chi Squared p value = 0.0006

	PR -ve	PR +ve	Total
HG	54	37	91
IG	23	33	56
LG	0	10	10
Total	77	80	157

PR and Histological Architecture

Chi squared p value = 0.0011

	PR -ve	PR +ve	Total
Crib	22	43	65
Micropap	12	3	15
Pap	4	8	12
Solid	39	26	65
Total	77	80	157

PR and Necrosis

Chi squared p value 0.0013

	PR -ve	PR +ve	Total
Marked	28	13	41
Mild	20	24	44
Moderate	21	17	38
None	8	26	34
Total	77	80	157

PR and Chronic Inflammation

Chi squared p value <0.0001

	PR -ve	PR +ve	Total
Marked	30	6	36
Mild	17	38	55
Moderate	17	17	34
None	12	18	30
Total	76	79	155

5.20 Appendix 6: DNA Extraction and Purification Protocols

PROTOCOL TO EXTRACT DNA FROM PARAFFIN EMBEDDED TISSUE

DAY 1 - SLIDE PREPARATION

8µm sections (number depending on size of specimen,) cut and mounted on slides.

Air dry overnight

DAY 2 -PARAFFIN REMOVAL & STAINING

Xylene 1 for 5'

Xylene 2 for 5'

100% ETOH 1 for 30"

100% ETOH 2 for 30"

70% ETOH for 30"

Wash in autoclaved water for 2'

Repeat wash with fresh autoclaved water

Stain with nuclear fast red (NFR) 1% - 4'

Wash with autoclaved water for 2'

Repeat wash with fresh autoclaved water

DAY 2 -MICRODISSECTION

Use H&E sections previously prepared and marked by Consultant Pathologists as a guide

Examine NFR stained sections under light microscope and using 12G sterile needle/scalpel scrape away unwanted areas and blow off with air spray

Incubate slides in 1M Na Thiocyanate at 37°C overnight.

DAY 3 – COMPLETE MICRODISSECTION AND TISSUE LYSIS

Wash in PBS-A, 10min on shaker, replace PBS then another 10mins on shaker.

Air dry.

Label appropriately (top and side) as many 1.5 ml eppendorf tubes as the samples you have (N) and add 180 µl of DNeasy Buffer ATL to each tube

Pipette 2-5 µl of ATL buffer at the edge of each section to moisten the tissue and lift using pipette tip or needle

Transfer the tissue in the appropriately labelled tube containing the ATL buffer.

Repeat steps 3-5 for all sections

Once microdissection is completed, add 20 µl Proteinase K, vortex for 15 seconds and spin.

Seal tubes with parafilm and incubate @56-58°C in a rotisserie/or on shaker in oven overnight.

Day 4 - DNA EXTRACTION

Reagents from Qiagen DNeasy Blood and Tissue Kit

Protocol: Purification of Total DNA from Animal Tissues (Spin-Column Protocol)

At 12 hours check that tissue fully lysed. If so, vortex well for 15 seconds and continue with DNeasy protocol by adding 200ul Buffer AL etc.

If tissue not fully lysed, add extra 20ul of DNeasy Proteinase K, vortex for 15 seconds and return to oven for 1-4hours, until tissue fully lysed

Once fully lysed continue with DNeasy protocol.

NB – Once lysed, samples can remain in buffer at -20 C until extraction proceeds

- Store at -20C after extraction

DCIS DNA Extraction using Qiagen FFPE Kit.

This protocol is adapted from the “QIAamp DNA FFPE Tissue handbook” October 2007. The original protocol is “Isolation of Genomic DNA from FFPE tissue sections” on page 15. This protocol is required AFTER manual microdissection of tissue sections on uncoated slides and follows incubation in 1M sodium thiocyanate prior to proteinase k treatment. It starts at point 14 of the original protocol. The reagents are all (except ETOH) part of the FFPE DNA extraction kit.

Between six and twelve samples should be prepared prior to extracting and be placed in 1.5ml eppendorfs tubes. Place frozen tube on wet ice and proceed with protocol.

In an eppendorf holder set out one row of samples followed by 1 set of labelled eluting columns on collecting tubes. Next 2 sets of collecting tubes followed by 2 sets of eppendorfs with labels corresponding to the sample numbers. The final set should be identified in some way as a spare.

Vortex all samples for approx. 15seconds

Add 200µl of buffer AL to each sample, if sample had extra proteinase k then add an additional 10-20 µl. The solution is viscous so pipette slowly. Re-vortex samples.

Add 200µl of 100% alcohol (ETOH) to each sample vortexing each sample immediately. The solution may start to precipitate, if this occurs keep vortexing until an almost clear homogenate is seen.

NB AL and ETOH can be premixed for all samples before adding.

Transfer the entire sample into a corresponding eluting column.

Centrifuge at 8000rpm (6000g) for one minute.

Place each eluting column into a new collection tube and discard the old one and contents.

Add 500µl of AW1 (check solution is correctly made up and ensure there is no precipitate present) to each sample. Spin at 8000rpm (6000g) for one minute. Put eluting column into new collection tube and discard old collecting tube and contents.

Add 500µl of AW2 and spin for 3.5 minutes at 13,500 rpm. Remove eluting and collection tubes ensuring that no liquid touches eluting tube membrane. If this occurs then re-spin.

Place eluting column into first labelled eppendorfs (not spare) and discard collecting tube and contents ensuring eluting tube membranes stay dry.

Add 20µl of ATE to eluting column making sure it is placed directly onto the membrane. Do this for all samples and then spin at 8000rpm for 1 minute.

Repeat step twelve spinning into same eppendorf to give total volume of ATE/DNA to 35-40µl.

Place eluting column into second eppendorf (spare) making sure lid is closed on first samples. Add 50µl of AE and spin at 8000rpm for one minute. Remove from centrifuge and discard eluting column and close sample lids.

Test for DNA quality using nanodrop and MultiplexPCR. If yield sufficient proceed to pico green testing and then store samples at -20°C. NB Nandrop deemed to be poor for quality analysis proceed straight to PicoGreen if desired.

5.21 Appendix 7. Qubit DNA quantities Quality: DNA Quibit Readings for MIP Array Analysis

Affymetrix ID	Grid Reference	Specimen type	Sample 1 Volume (ul)	Qubit Reading (ug/μl)	Qubit stock conc	Total DNA per original sample (ng per 40μl)
DCISJPB1OK 001	A1	Triple Negative Pure DCIS	40	0.456	91.2	3648
DCISJPB1OK 002	A2	Triple Negative Pure DCIS	40	1.030	206.0	8240
DCISJPB1OK 003	A3	Triple Negative Pure DCIS	40	0.102	20.4	816
DCISJPB1OK 004	A4	Triple negative DCIS associated with tumour	40	0.0479	9.58	383.2
DCISJPB1OK 005	A5	Triple negative Tumour associated with DCIS	40	1.15	230	9200
DCISJPB1OK 006	A6	Triple Negative Pure DCIS	40	0.188	37.6	1504
DCISJPB1OK 007	A7	HER2 Positive	40	0.91	18.2	728
DCISJPB1OK 008	A8	HER2 Positive	40	1.39	27.8	1112
DCISJPB1OK 009	A9	Pure DCIS ER positive	40	1.600	320	12800
DCISJPB1OK 010	A10	Triple negative DCIS associated with tumour	40	0.668	134	5360
DCISJPB1OK 011	A11	Triple negative Tumour associated with DCIS	40	0.917	183	7320
DCISJPB1OK 012	B1	Matched Normal from Triple negative DCIS and Tumour	40	0.818	16.36	654.4
DCISJPB1OK 013	B2	Triple negative DCIS associated with tumour	40	0.858	172	6880

DCISJPB1OK 014	B3	Triple negative Tumour associated with DCIS	40	1.13	226	9040
----------------	----	---	----	------	-----	------

Affymetrix ID	Grid Reference	Specimen type	Sample 1 Volume (ul)	Qubit Reading (ug/μl)	Qubit stock conc	Total DNA per original sample (ng per 40μl)
DCISJPB1OK 015	B4	Matched Normal from Triple negative DCIS and Tumour	40	1.69	33.8	1352
DCISJPB1OK 016	B5	HER2 Positive	40	0.778	15.56	622.4
DCISJPB1OK 017	B6	Pure DCIS ER positive	40	0.643	128.6	5144
DCISJPB1OK 018	B7	Triple negative DCIS associated with tumour	40	1.24	248	9920
DCISJPB1OK 019	B8	Triple negative Tumour associated with DCIS	40	0.831	166	6640
DCISJPB1OK 020	B9	Triple Negative Pure DCIS	40	0.222	44.4	1776
DCISJPB1OK 021	B10	Pure DCIS ER positive	40	2.510	502	20080
DCISJPB1OK 022	B11	Triple Negative Pure DCIS	40	0.078	15.7	627.2
DCISJPB1OK 023	B12	HER2 Positive	40	1.06	21.2	848
DCISJPB1OK 024	C1	Pure DCIS ER positive	40	1.670	334	13360
DCISJPB1OK 025	C2	HER2 Positive	40	1.46	29.2	1168

DCISJPB1OK 026	C3	Triple negative DCIS associated with tumour	40	0.782	156	6240
DCISJPB1OK 027	C4	Triple negative Tumour associated with DCIS	40	0.936	187	7480
DCISJPB1OK 028	C5	Matched Normal from Triple negative DCIS and Tumour	40	1.59	31.8	1272

Affymetrix ID	Grid Reference	Specimen type	Sample 1 Volume (ul)	Qubit Reading (ug/μl)	Qubit stock conc	Total DNA per original sample (ng per 40μl)
DCISJPB1OK 029	C6	Triple negative DCIS associated with tumour	40	1.59	318	12720
DCISJPB1OK 030	C7	Triple negative Tumour associated with DCIS	40	1.37	274	10960
DCISJPB1OK 031	C8	Matched Normal from Triple negative DCIS and Tumour	40	0.297	5.94	237.6
DCISJPB1OK 032	C9	Triple Negative Pure DCIS	40	0.153	30.6	1224
DCISJPB1OK 033	C10	Triple negative DCIS associated with tumour	40	1.68	336	13440
DCISJPB1OK 034	C11	Triple negative Tumour associated with DCIS	40	2.68	536	21440
DCISJPB1OK 035	D1	Matched Normal from Triple negative DCIS and Tumour	40	1.84	36.8	1472
DCISJPB1OK 036	D2	HER2 Positive	40	0.951	19.02	760.8
DCISJPB1OK 037	D3	HER2 Positive	40	1.75	35	1400

DCISJPB1OK 038	D4	Triple Negative Pure DCIS	40	0.049	9.7	389.6
DCISJPB1OK 039	D5	Matched Normal from Triple negative DCIS and Tumour	40	1.28	25.6	1024
DCISJPB1OK 040	D6	Pure DCIS ER positive	40	0.447	89.4	3576
DCISJPB1OK 041	D7	Triple negative DCIS associated with tumour	40	0.265	53	2120
DCISJPB1OK 042	D8	Triple negative Tumour associated with DCIS	40	1.86	372	14880

Affymetrix ID	Grid Reference	Specimen type	Sample 1 Volume (ul)	Qubit Reading (ug/μl)	Qubit stock conc	Total DNA per original sample (ng per 40μl)
DCISJPB1OK 043	D9	Triple negative DCIS associated with tumour	40	0.647	129	5160
DCISJPB1OK 044	D10	Triple negative Tumour associated with DCIS	40	1.23	246	9840
DCISJPB1OK 045	D11	Matched Normal from Triple negative DCIS and Tumour	40	1.86	372	14880
DCISJPB1OK 046	D12	Triple negative DCIS associated with tumour	40	1.13	226	9040
DCISJPB1OK 047	E1	Triple negative Tumour associated with DCIS	40	2.12	434	17360
DCISJPB1OK 048	E2	Matched Normal from Triple negative DCIS and Tumour	40	0.558	11.16	446.4
DCISJPB1OK 049	E3	Pure DCIS ER positive	40	2.330	466	18640

DCISJPB1OK 050	E4	Pure DCIS ER positive	40	1.150	230	9200
DCISJPB1OK 051	E5	DCIS ER positive associated with tumour	40	0.702	140.4	5616
DCISJPB1OK 052	E6	Tumour ER positive associated with DCIS	40	0.780	156	6240
DCISJPB1OK 053	E7	Matched Normal from ER positive DCIS and Tumour	40	0.582	11.64	465.6
DCISJPB1OK 054	E8	DCIS ER positive associated with tumour	40	0.105	21	840
DCISJPB1OK 055	E9	Tumour ER positive associated with DCIS	40	1.190	238	9520

Affymetrix ID	Grid Reference	Specimen type	Sample 1 Volume (ul)	Qubit Reading (ug/μl)	Qubit stock conc	Total DNA per original sample (ng per 40μl)
DCISJPB1OK 056	E10	Matched Normal from ER positive DCIS and Tumour	40	2.06	41.2	1648
DCISJPB1OK 057	E11	DCIS ER positive associated with tumour	40	0.611	122.2	4888
DCISJPB1OK 058	F1	Tumour ER positive associated with DCIS	40	0.788	157.6	6304
DCISJPB1OK 059	F2	Matched Normal from ER positive DCIS and Tumour	40	0.0883	1.766	70.64
DCISJPB1OK 060	F3	DCIS ER positive associated with tumour	40	0.367	73.4	2936

DCISJPB1OK 061	F4	Tumour ER positive associated with DCIS	40	1.760	352	14080
DCISJPB1OK 062	F5	Matched Normal from ER positive DCIS and Tumour	40	1.46	29.2	1168
DCISJPB1OK 063	F6	DCIS ER positive associated with tumour	40	1.150	230	9200
DCISJPB1OK 064	F7	Tumour ER positive associated with DCIS	40	1.350	270	10800
DCISJPB1OK 065	F8	Pure DCIS ER positive	40	0.277	55.4	2216
DCISJPB1OK 066	F9	Pure DCIS ER positive	40	1.560	312	12480
DCISJPB1OK 067	F10	DCIS ER positive associated with tumour	40	0.568	113.6	4544
DCISJPB1OK 068	F11	Tumour ER positive associated with DCIS	40	0.794	158.8	6352
DCISJPB1OK 070	G1	Matched Normal from ER positive DCIS and Tumour	40	1.29	25.8	1032

Affymetrix ID	Grid Reference	Specimen type	Sample 1 Volume (ul)	Qubit Reading (ug/μl)	Qubit stock conc	Total DNA per original sample (ng per 40μl)
DCISJPB1OK 069	F12	HER2 Positive	40	2.19	43.8	1752
DCISJPB1OK 071	G2	Pure DCIS ER positive	40	0.571	114.2	4568
DCISJPB1OK 072	G3	Triple Negative Pure DCIS	40	0.062	12.4	494.4

DCISJPB1OK 073	G4	Triple Negative Pure DCIS	40	0.400	80.0	3200
DCISJPB1OK 074	G5	Triple Negative Pure DCIS	40	0.835	167.0	6680
DCISJPB1OK 075	G6	Triple Negative Pure DCIS	40	0.085	17.0	680
DCISJPB1OK 076	G7	DCIS ER positive associated with tumour	40	0.628	125.6	5024
DCISJPB1OK 077	G8	Tumour ER positive associated with DCIS	40	0.234	46.8	1872
DCISJPB1OK 078	G9	DCIS ER positive associated with tumour	40	0.095	19.08	763.2
DCISJPB1OK 079	G10	Tumour ER positive associated with DCIS	40	0.189	37.8	1512
DCISJPB1OK 080	G11	Matched Normal from ER positive DCIS and Tumour	40	1.03	20.6	824
DCISJPB1OK 081	H1	DCIS ER positive associated with tumour	40	0.932	186.4	7456
DCISJPB1OK 082	H2	Tumour ER positive associated with DCIS	40	0.875	175	7000
DCISJPB1OK 083	H3	Matched Normal from ER positive DCIS and Tumour	40	2.79	55.8	2232

5.22 Appendix 8. MIP Array Analysis- DNA Samples sent to Affymetrix

List of specimens

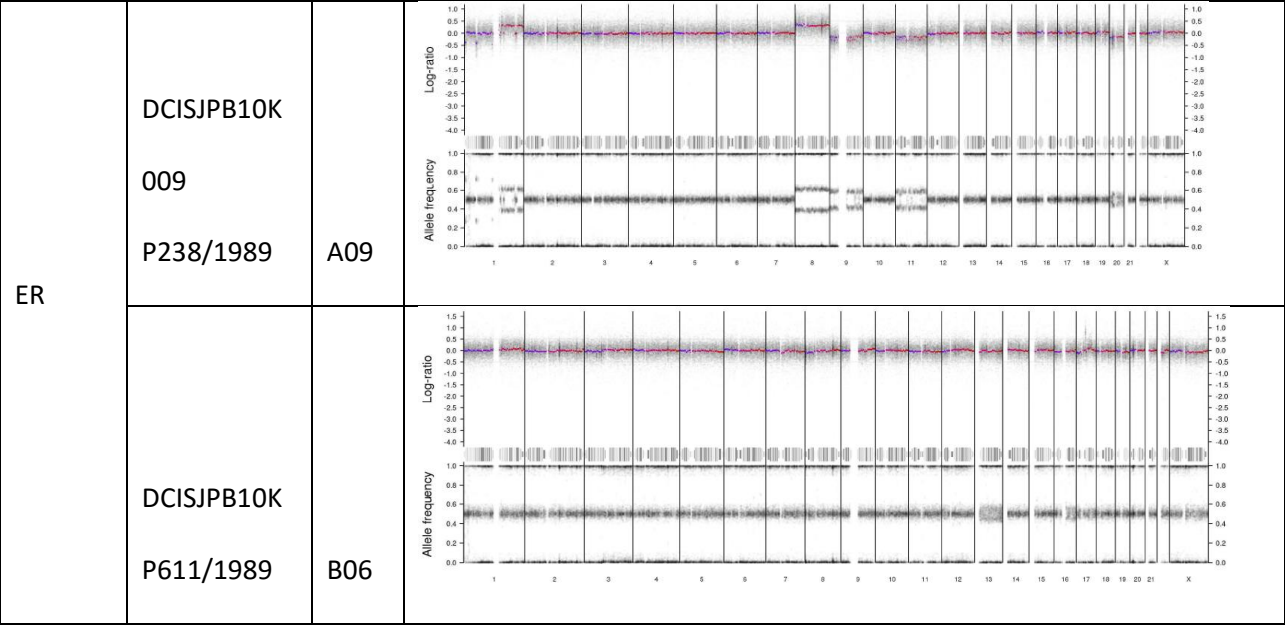
UNIQUE IDENTIFIER	PLATE REFERENCE	Specimen type
DCISJPB1OK 001	A1	Triple Negative Pure DCIS
DCISJPB1OK 002	A2	Triple Negative Pure DCIS
DCISJPB1OK 003	A3	Triple Negative Pure DCIS
DCISJPB1OK 004	A4	Triple negative DCIS
DCISJPB1OK 005	A5	Triple negative Tumour
DCISJPB1OK 006	A6	Triple Negative Pure DCIS
DCISJPB1OK 007	A7	HER2 Positive
DCISJPB1OK 008	A8	HER2 Positive
DCISJPB1OK 009	A9	Pure DCIS ER positive
DCISJPB1OK 010	A10	Triple negative DCIS
DCISJPB1OK 011	A11	Triple negative Tumour
DCISJPB1OK 012	B1	Matched Normal from Triple negative DCIS and Tumour
DCISJPB1OK 013	B2	Triple negative DCIS
DCISJPB1OK 014	B3	Triple negative Tumour
DCISJPB1OK 015	B4	Matched Normal from Triple negative DCIS and Tumour
DCISJPB1OK 016	B5	HER2 Positive
DCISJPB1OK 017	B6	Pure DCIS ER positive
DCISJPB1OK 018	B7	Triple negative DCIS
DCISJPB1OK 019	B8	Triple negative Tumour
DCISJPB1OK 020	B9	Triple Negative Pure DCIS
DCISJPB1OK 021	B10	Pure DCIS ER positive
DCISJPB1OK 022	B11	Triple Negative Pure DCIS
DCISJPB1OK 023	B12	HER2 Positive

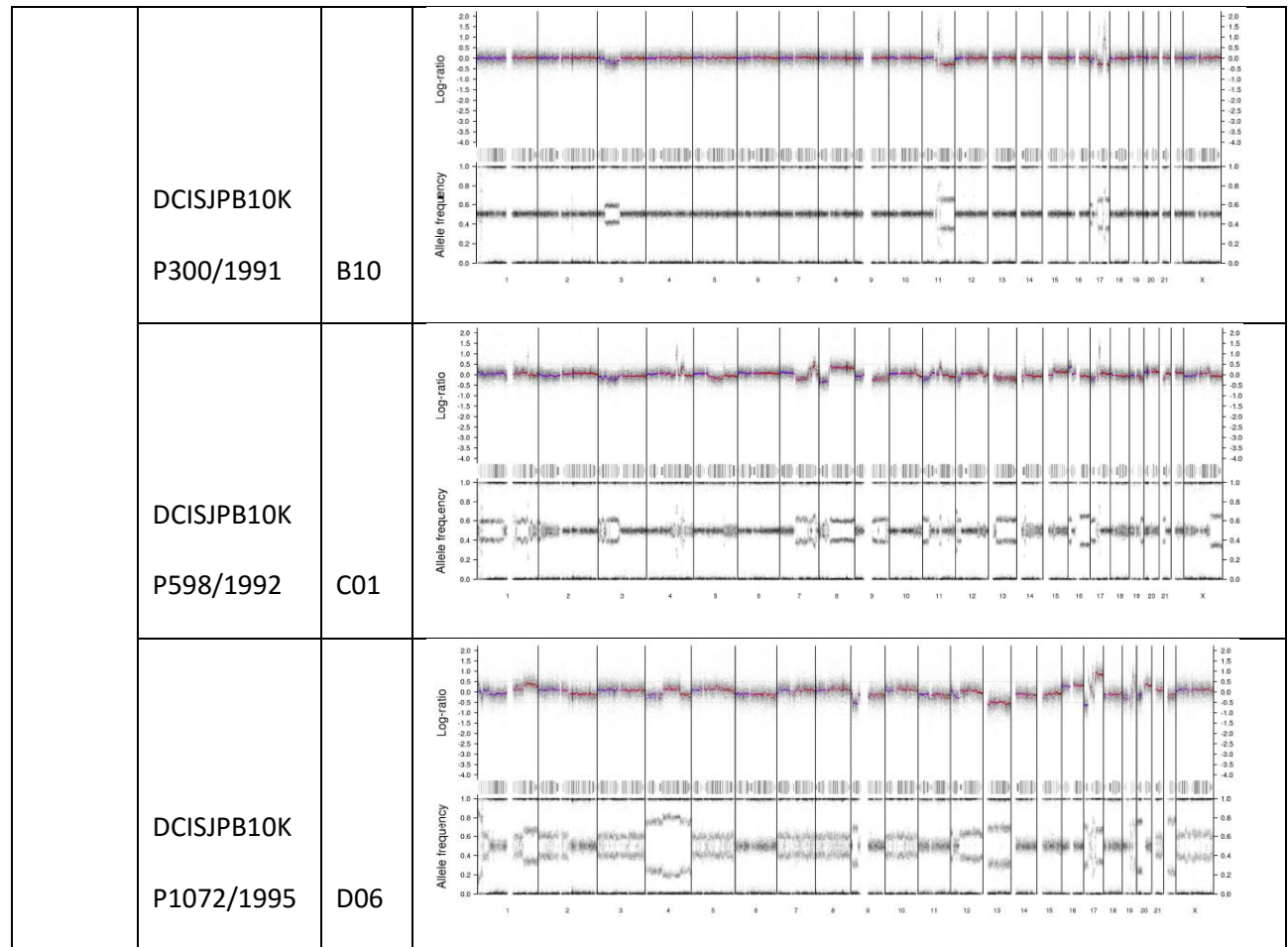
DCISJPB1OK 024	C1	Pure DCIS ER positive
DCISJPB1OK 025	C2	HER2 Positive
DCISJPB1OK 026	C3	Triple negative DCIS
DCISJPB1OK 027	C4	Triple negative Tumour
DCISJPB1OK 028	C5	Matched Normal from Triple negative DCIS and Tumour
DCISJPB1OK 029	C6	Triple negative DCIS
DCISJPB1OK 030	C7	Triple negative Tumour
DCISJPB1OK 031	C8	Matched Normal from Triple negative DCIS and Tumour
DCISJPB1OK 032	C9	Triple Negative Pure DCIS
DCISJPB1OK 033	C10	Triple negative DCIS
DCISJPB1OK 034	C11	Triple negative Tumour
DCISJPB1OK 035	D1	Matched Normal from Triple negative DCIS and Tumour
DCISJPB1OK 036	D2	HER2 Positive
DCISJPB1OK 037	D3	HER2 Positive
DCISJPB1OK 038	D4	Triple Negative Pure DCIS
DCISJPB1OK 039	D5	Matched Normal from Triple negative DCIS and Tumour
DCISJPB1OK 040	D6	Pure DCIS ER positive
DCISJPB1OK 041	D7	Triple negative DCIS
DCISJPB1OK 042	D8	Triple negative Tumour
DCISJPB1OK 043	D9	Triple negative DCIS
DCISJPB1OK 044	D10	Triple negative Tumour
DCISJPB1OK 045	D11	Matched Normal from Triple negative DCIS and Tumour
DCISJPB1OK 046	D12	Triple negative DCIS
DCISJPB1OK 047	E1	Triple negative Tumour
DCISJPB1OK 048	E2	Matched Normal from Triple negative DCIS and Tumour

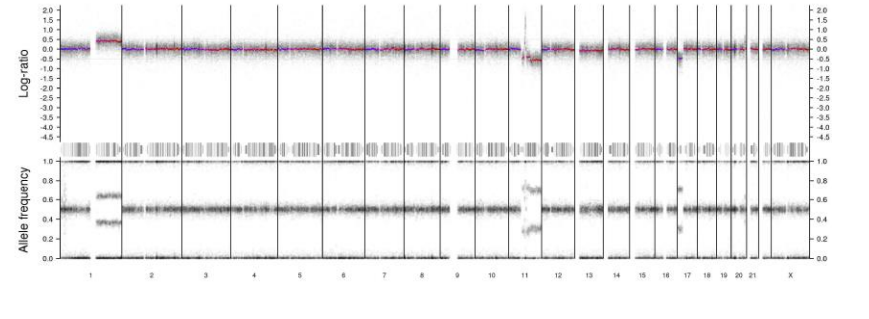
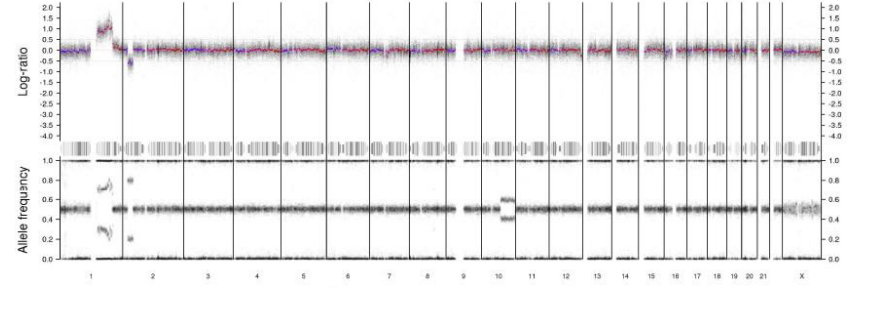
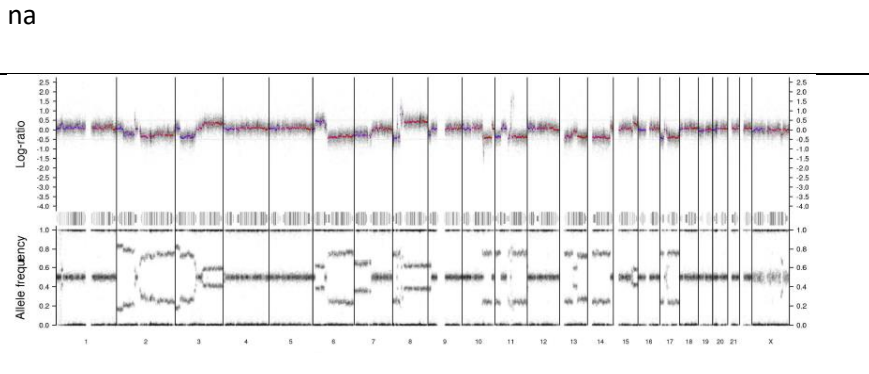
DCISJPB1OK 049	E3	Pure DCIS ER positive
DCISJPB1OK 050	E4	Pure DCIS ER positive
DCISJPB1OK 051	E5	Tumour ER positive (basal negative)
DCISJPB1OK 052	E6	DCIS ER positive (basal negative)
DCISJPB1OK 053	E7	Matched Normal from ER positive DCIS and Tumour
DCISJPB1OK 054	E8	Tumour ER positive (basal negative)
DCISJPB1OK 055	E9	DCIS ER positive (basal negative)
DCISJPB1OK 056	E10	Matched Normal from ER positive DCIS and Tumour
DCISJPB1OK 057	E11	Tumour ER positive (basal negative)
DCISJPB1OK 058	F1	DCIS ER positive (basal negative)
DCISJPB1OK 059	F2	Matched Normal from ER positive DCIS and Tumour
DCISJPB1OK 060	F3	Tumour ER positive (basal negative)
DCISJPB1OK 061	F4	DCIS ER positive (basal negative)
DCISJPB1OK 062	F5	Matched Normal from ER positive DCIS and Tumour
DCISJPB1OK 063	F6	Tumour ER positive (basal negative)
DCISJPB1OK 064	F7	DCIS ER positive (basal negative)
DCISJPB1OK 065	F8	Pure DCIS ER positive
DCISJPB1OK 066	F9	Pure DCIS ER positive
DCISJPB1OK 067	F10	Tumour ER positive (basal negative)
DCISJPB1OK 068	F11	DCIS ER positive (basal negative)
DCISJPB1OK 069	F12	Matched Normal from ER positive DCIS and Tumour
DCISJPB1OK 070	G1	HER2 Positive

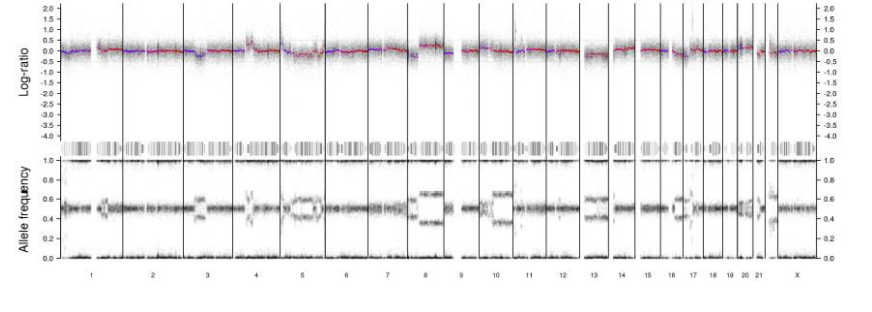
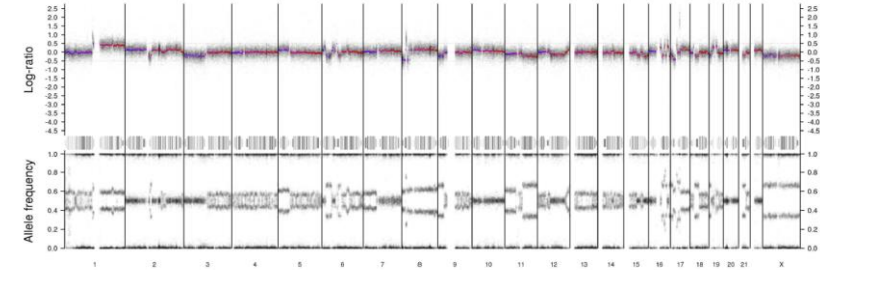
DCISJPB1OK 071	G2	Triple Negative Pure DCIS
DCISJPB1OK 072	G3	Triple Negative Pure DCIS
DCISJPB1OK 073	G4	Triple Negative Pure DCIS
DCISJPB1OK 074	G5	Triple Negative Pure DCIS
DCISJPB1OK 075	G6	Pure DCIS ER positive
DCISJPB1OK 076	G7	Tumour ER positive (basal negative)
DCISJPB1OK 077	G8	DCIS ER positive (basal negative)
DCISJPB1OK 078	G9	Tumour ER positive (basal negative)
DCISJPB1OK 079	G10	DCIS ER positive (basal negative)
DCISJPB1OK 080	G11	Matched Normal from ER positive DCIS and Tumour
DCISJPB1OK 081	H1	Tumour ER positive (basal negative)
DCISJPB1OK 082	H2	DCIS ER positive (basal negative)
DCISJPB1OK 083	H3	Matched Normal from ER positive DCIS and Tumour

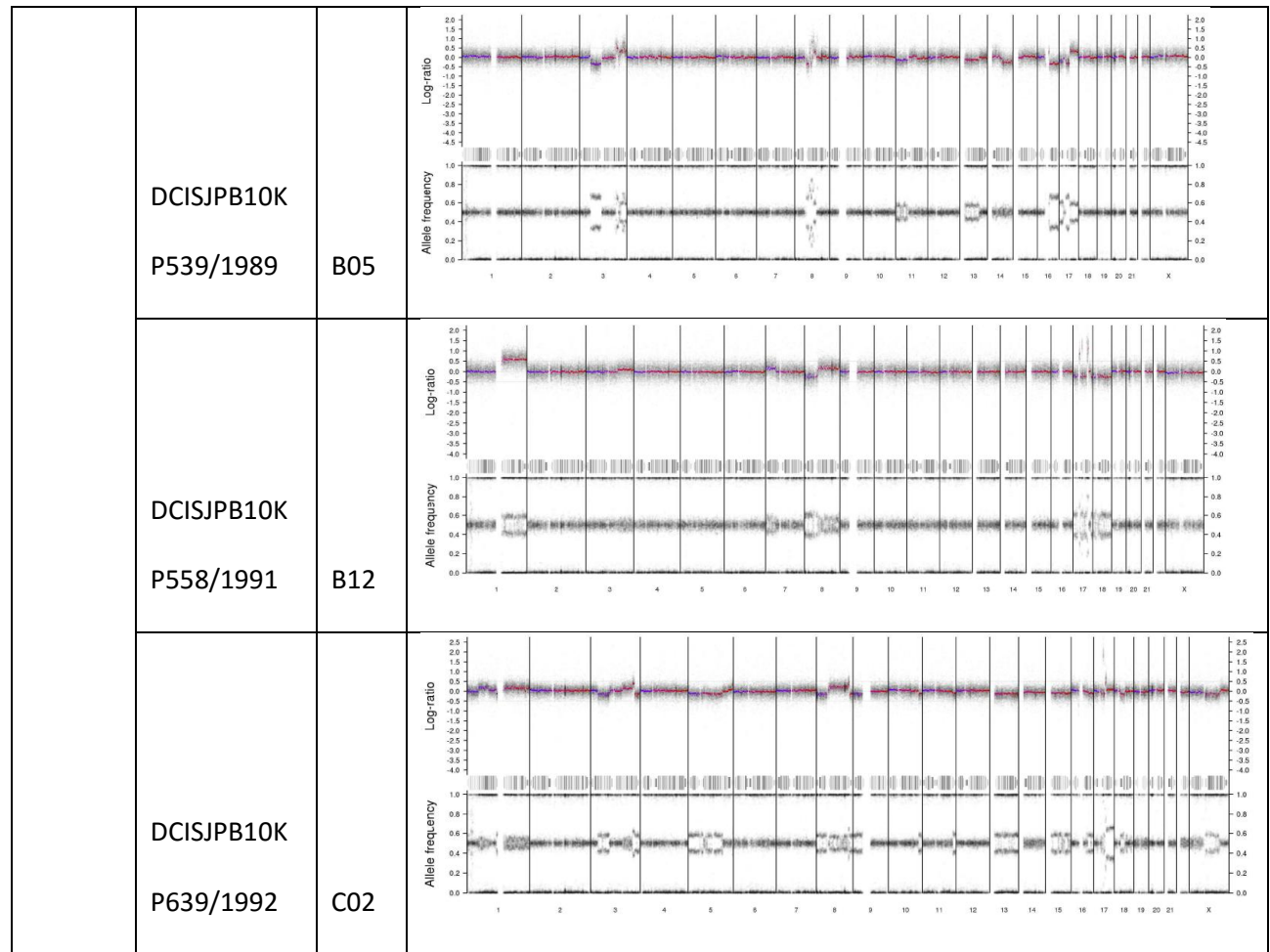
5.23 Appendix 9 MIP Array Charts

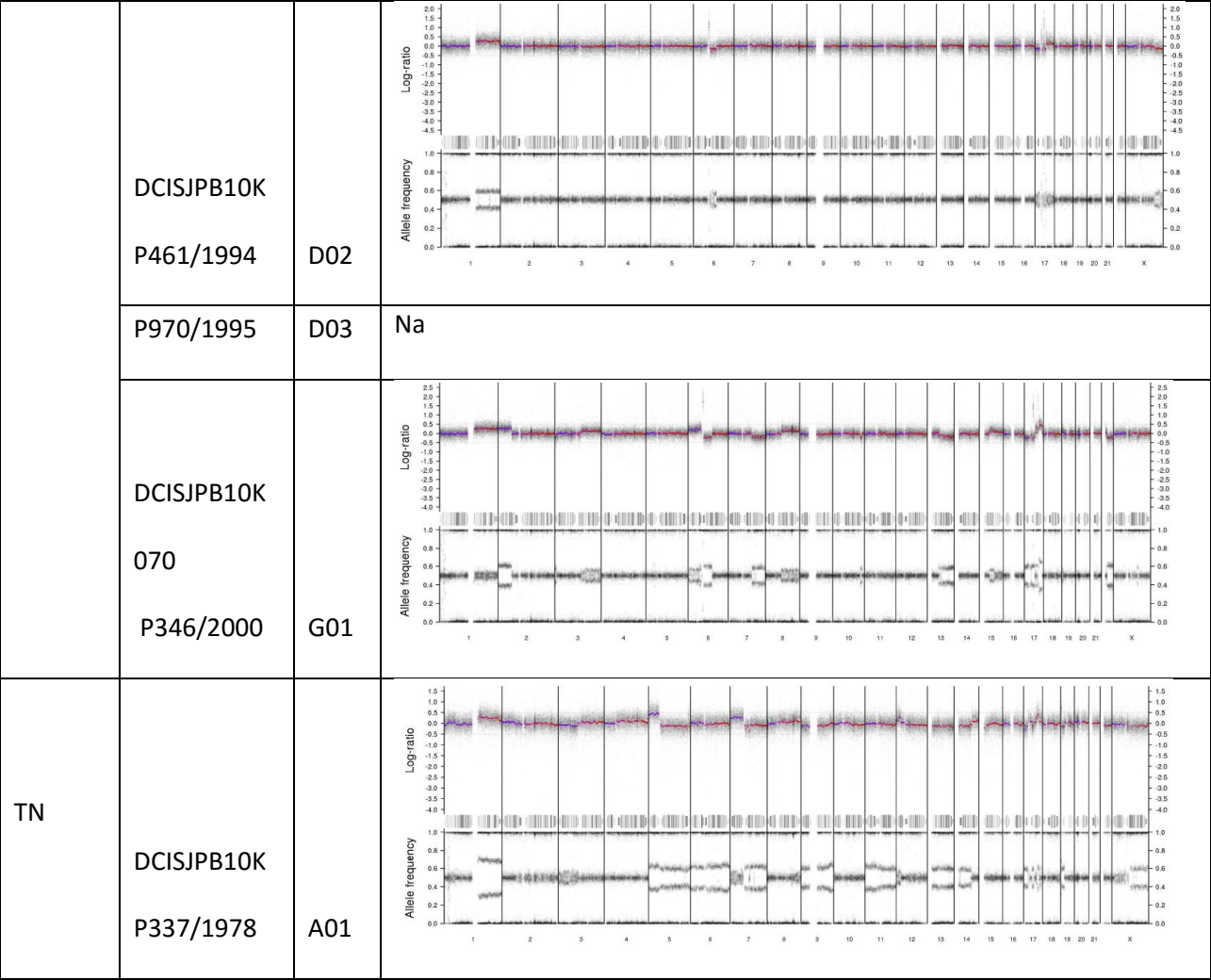


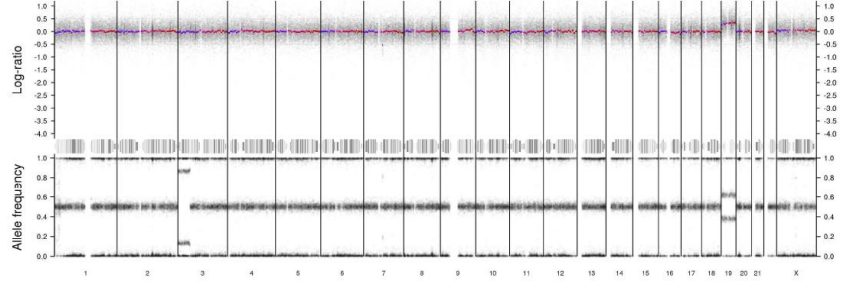
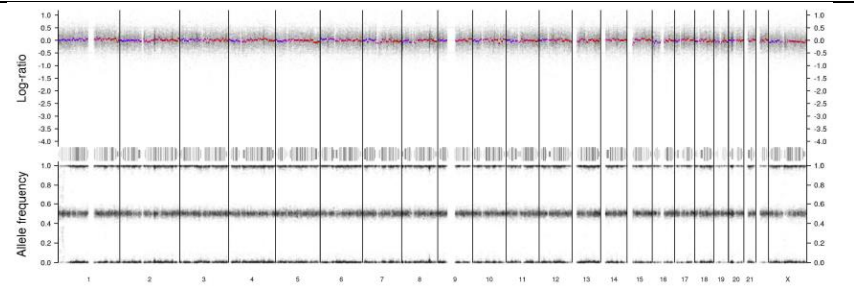


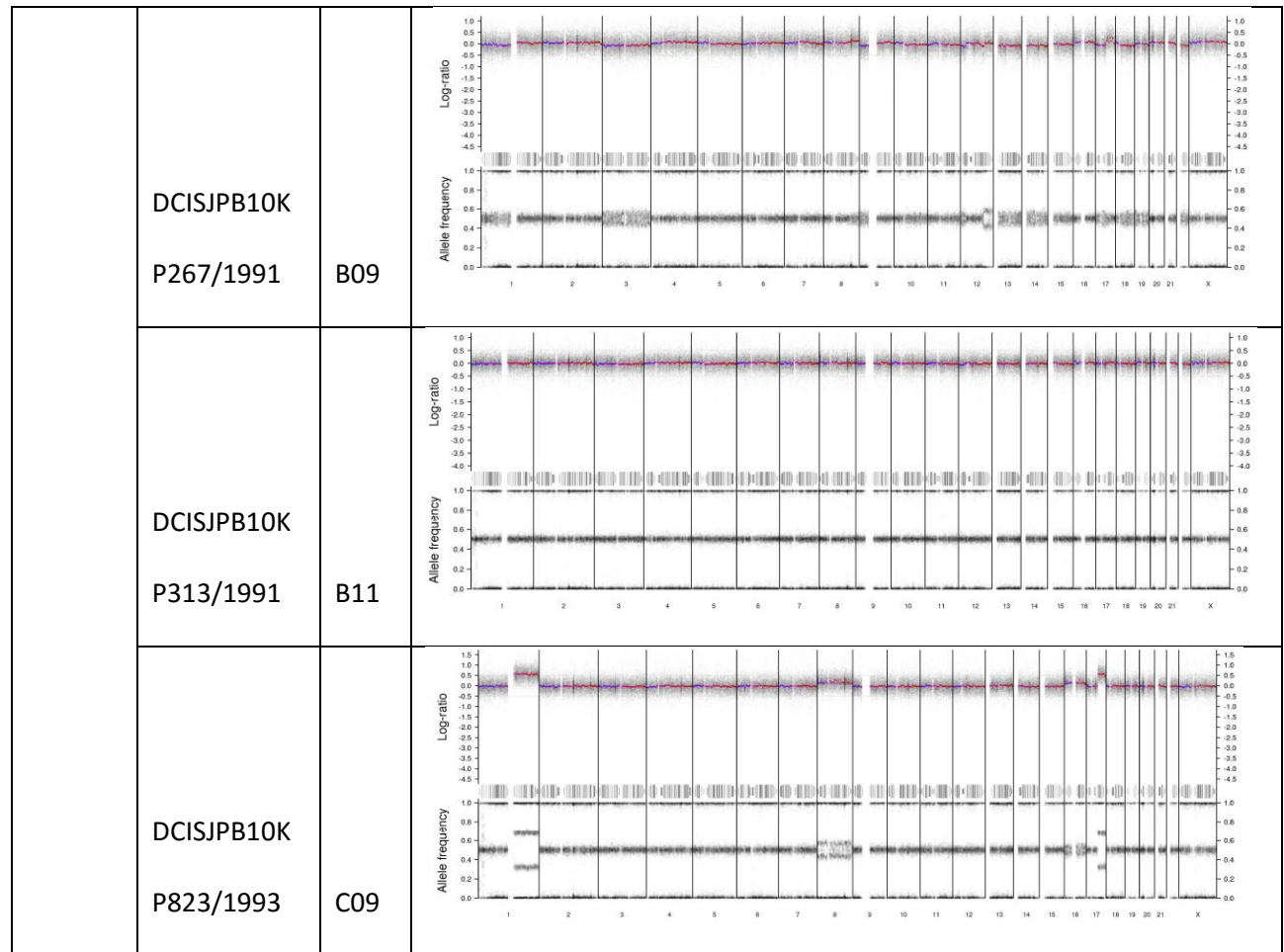
	DCISJPB10K P413/1997	E03	 <p>Genomic plots for E03. The top panel shows Log-ratio (y-axis: -4.5 to 2.0) across chromosomes 1-21 and X. The bottom panel shows Allele frequency (y-axis: 0.0 to 1.0) across the same chromosomes. The plots show a relatively stable baseline with minor fluctuations.</p>
	DCISJPB10K P754/1997	E04	 <p>Genomic plots for E04. The top panel shows Log-ratio (y-axis: -4.0 to 2.0) across chromosomes 1-21 and X. The bottom panel shows Allele frequency (y-axis: 0.0 to 1.0) across the same chromosomes. The plots show a relatively stable baseline with minor fluctuations.</p>
	P198/1999	F08	na
	DCISJPB10K 66 P1194/1999	F09	 <p>Genomic plots for F09. The top panel shows Log-ratio (y-axis: -4.0 to 2.5) across chromosomes 1-21 and X. The bottom panel shows Allele frequency (y-axis: 0.0 to 1.0) across the same chromosomes. The plots show a relatively stable baseline with minor fluctuations.</p>

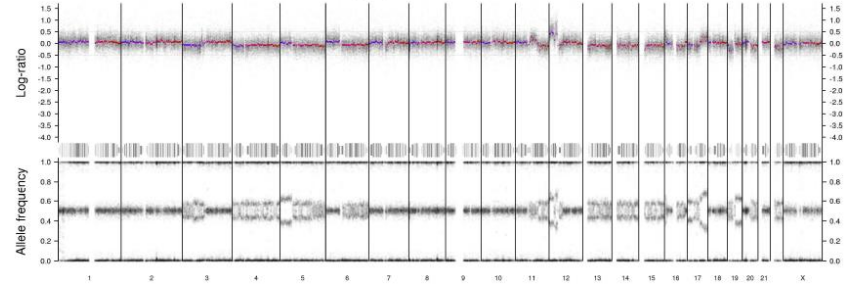
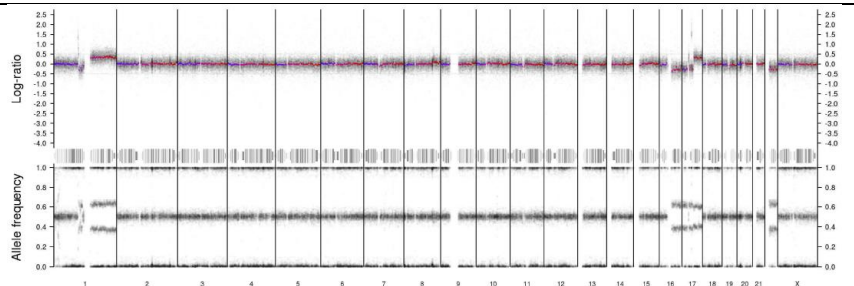
	DCISJPB10K P430/2000	G02	na
HER2	DCISJPB10K P914/1988	A07	 <p>Genomic plots for sample P914/1988. The top plot shows Log-ratio (y-axis: -4.0 to 2.0) across chromosomes 1-21 and X. The bottom plot shows Allele frequency (y-axis: 0.0 to 1.0) across the same chromosomes. The plots show a relatively stable Log-ratio around 0 and Allele frequency around 0.5 across most chromosomes, with some minor fluctuations.</p>
	DCISJPB10K P10/1989	A08	 <p>Genomic plots for sample P10/1989. The top plot shows Log-ratio (y-axis: -4.0 to 2.0) across chromosomes 1-21 and X. The bottom plot shows Allele frequency (y-axis: 0.0 to 1.0) across the same chromosomes. The plots show a relatively stable Log-ratio around 0 and Allele frequency around 0.5 across most chromosomes, with some minor fluctuations.</p>

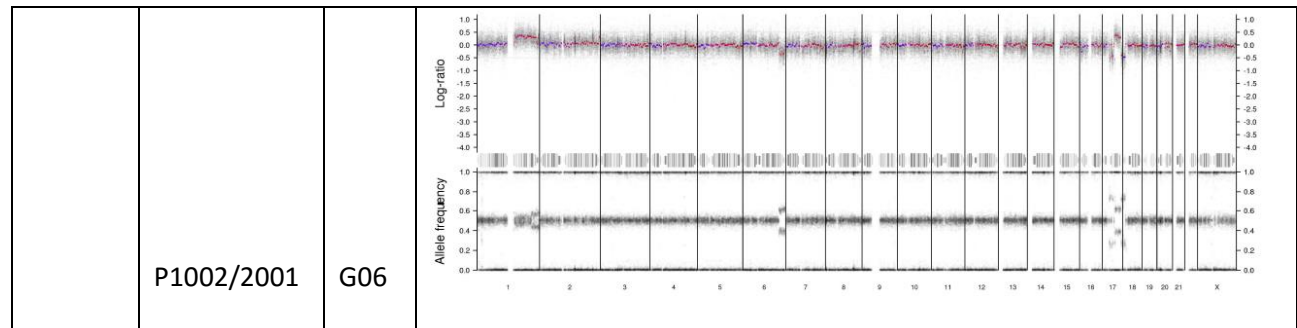


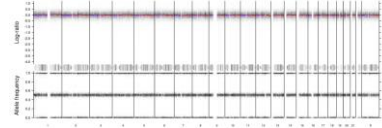
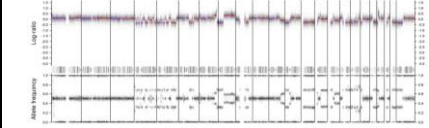
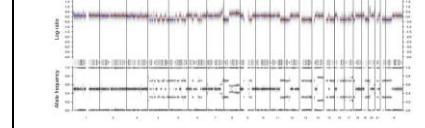
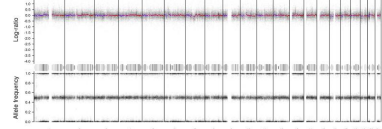
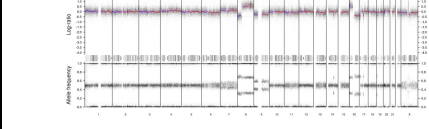
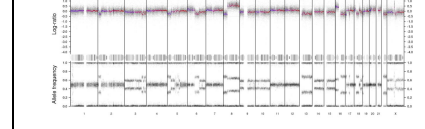
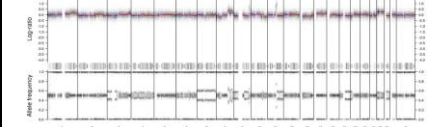
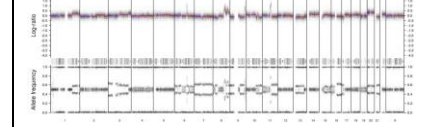
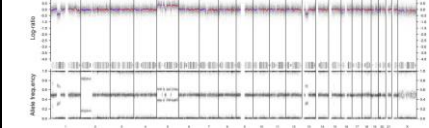

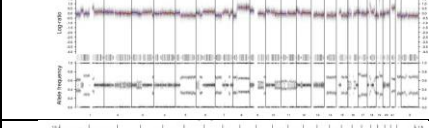
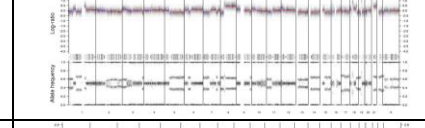
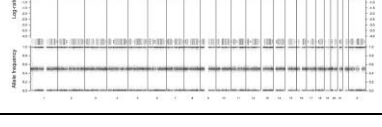
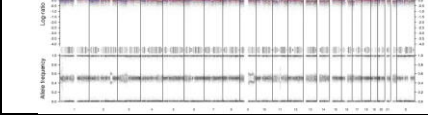
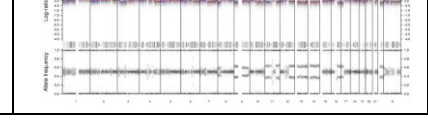


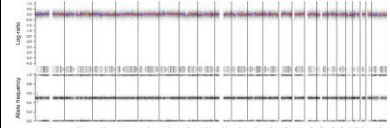
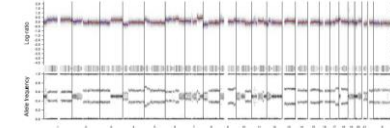
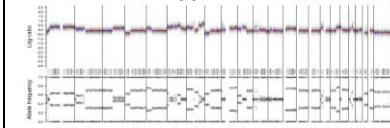
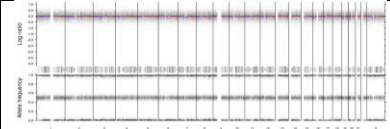
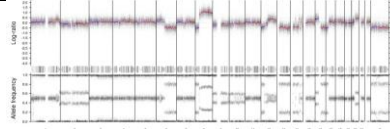
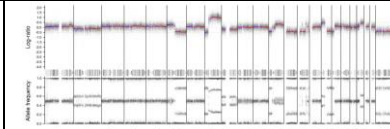
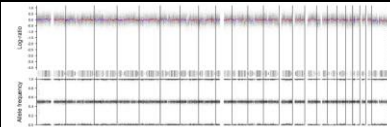
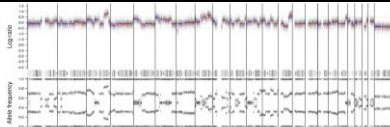
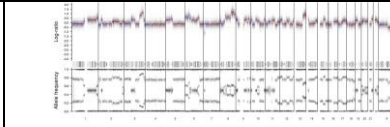
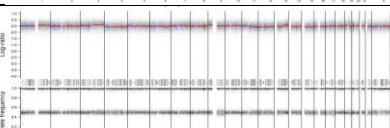
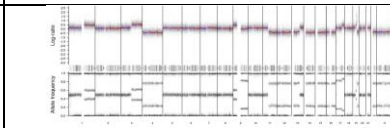
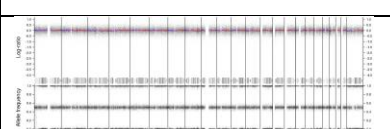
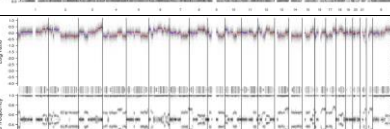
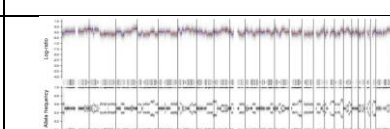

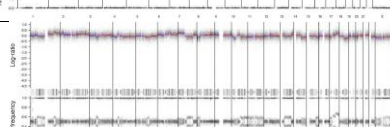
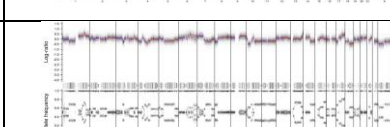
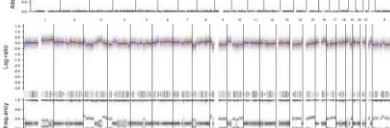
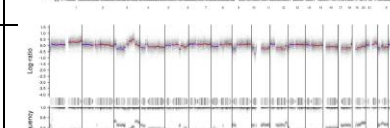
	DCISJPB10K P997/1982	A02	 <p>Genomic plots for sample P997/1982. The top panel shows Log-ratio (y-axis: -4.0 to 1.0) across chromosomes 1-21 and X. The bottom panel shows Allele frequency (y-axis: 0.0 to 1.0) across the same chromosomes. The plots show a relatively flat Log-ratio profile and stable allele frequencies across most chromosomes, with some minor fluctuations on chromosomes 3, 7, and 19.</p>
	DCISJPB10K P591/1984	A03	 <p>Genomic plots for sample P591/1984. The top panel shows Log-ratio (y-axis: -4.0 to 1.0) across chromosomes 1-21 and X. The bottom panel shows Allele frequency (y-axis: 0.0 to 1.0) across the same chromosomes. The plots show a relatively flat Log-ratio profile and stable allele frequencies across most chromosomes, with some minor fluctuations on chromosomes 3, 7, and 19.</p>
	DCISJPB10K P621/1988	A06	Na

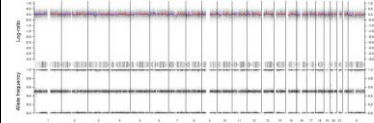
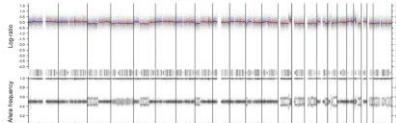
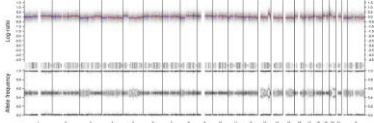
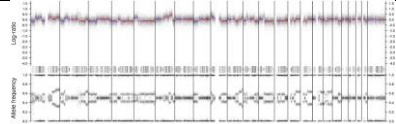
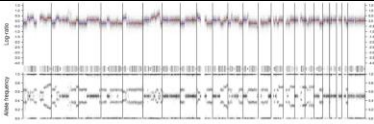
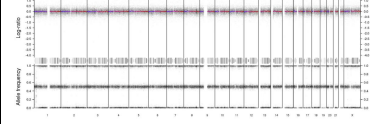
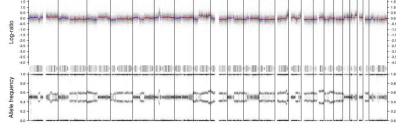
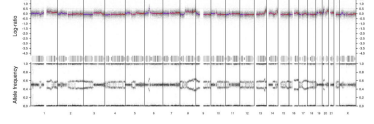
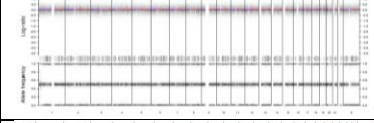
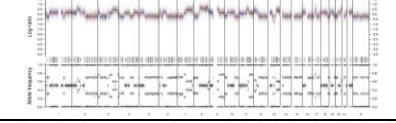
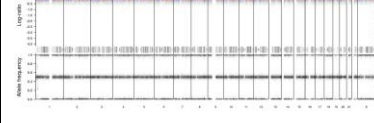
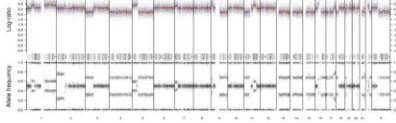
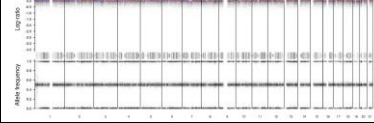
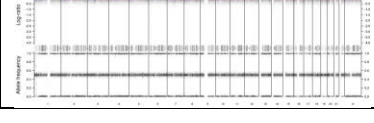


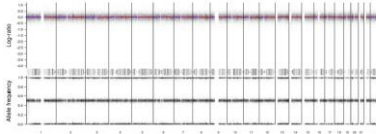
	DCISJPB10K P990/1995	D04	na
	DCISJPB10K P1074/2000	G03	 <p>Genomic plots for G03. The top panel shows Log-ratio (y-axis: -4.0 to 1.5) across chromosomes 1-21 and X. The bottom panel shows Allele frequency (y-axis: 0.0 to 1.0) across the same chromosomes. The plots show a relatively stable signal across the genome with some minor fluctuations.</p>
	DCISJPB10K P1093/2000	G04	na
	P701/2001	G05	 <p>Genomic plots for G05. The top panel shows Log-ratio (y-axis: -4.0 to 2.5) across chromosomes 1-21 and X. The bottom panel shows Allele frequency (y-axis: 0.0 to 1.0) across the same chromosomes. The plots show a relatively stable signal across the genome with some minor fluctuations.</p>



	case no	normal		DCIS		invasive breast disease	
ER	P123/1998	E07		E06		E05	
	P1310/2004	H03		H02		H01	
	P168/1999			F07		F06	
	P258/2004			G08		G07	
	P339/2000			F11		F10	
	P367/2004	G11		G10		G09	

Triple Negati ve	P46/1999	F05		F04		F03	
	P585/1998	E10		E09		E08	
	P769/1998	F02		F01		E11	
	P110/1987			A04		A05	
	P126/1997	D11		D09		D10	
	P137/1997	E02		D12		E01	
	P220/1990			B07		B08	

P456/1989	B01		A10		A11	
P497/1989			B02		B03	
P9/1994	D01		C10		C11	
P540/1993	C08		C06			
P730/1992	C05		C03			
P1012/1995	D05					
P400/2000	F12					

	P536/1989	B04		
	P1202/1995		D07 